A Variation Of Forms:
The Cognitive Neuropsychology Of Primary Progressive Aphasia

Chris Westbury
Department Of Psychology
McGill University, Montreal
June, 1995

A thesis submitted to the Faculty Of Graduate Studies and Research in partial fulfillment of the requirements of the degree of Ph.D.

ISBN 0-612-12509-2

"If we saw as much of the world
as we do not see, we should
be aware, in all probability,
of a perpetual multiplication
and variation of forms."

Michel De Montaigne
Essays

Abstract

Mesulam's (1982) report describing six patients with a slowly progressive aphasia without accompanying signs of dementia led to the recognition of a syndrome now known as Primary Progressive Aphasia (PPA). Many more patients have been described since Mesulam's description was published: 171 published contacts with 112 patients diagnosed with the syndrome are reviewed in this thesis. However, the published literature is both unsystematic and incomplete, making it difficult to place the findings into a coherent theoretical framework. In addition, no previous work has specifically attempted to specify the difference between PPA and dementia of Alzheimer's type (DAT), although the two disorders are easily confused since many language deficits can masquerade as memory or cognitive deficits. This thesis is an attempt to remedy these short-comings. Using a serial case study approach, the linguistic deficits of 11 PPA patients are analyzed in a cognitive neuropsychological framework, and contrasted with the linguistic deficits of a group of 11 DAT patients. Several tools were developed or refined specifically to manage the details of such an analysis. These include a production system designed to infer functional dissociations in the language system from an analysis of patient performance on a computerized version of the Psycholinguistic Assessment Battery (Caplan & Bub, 1990). Although the results suggest that there is great heterogeneity of symptoms within the PPA population, several deficits which may prove useful in making the differential diagnosis are examined closely, including deficits in reading, abstract word comprehension, auditory comprehension, affixed word processing, and semantic access. Implications of the findings for current models of ianguage processing are discussed

Résumé

Le rapport dans lequel Mesulam (1982) décrit six patients atteints d'aphasie progressive lente sans signe de démence a conduit à l'identification d'un syndrome connu sous le nom d'aphasie progressive primitive (APP). Depuis la parution du rapport de Mesulam, de nombreux autres patients ont fait l'objet d'une description: 171 contacts avec 112 patients chez qui ce syndrome a été diagnositiqué sont résumés dans cette thèse. Les publications sur ce sujet étant à la fois peu systématiques et incomplètes, il est toutefois difficile de placer les résultats dans un cadre théorique cohérent. Aucune recherche antérieure n'a par ailleurs cherché à préciser la différence entre l'APP et la démence de type Alzheimer (DTA) alors qu'il est facile de confondre les deux troubles puisque de nombreux déficits linguistiques peuvent passer pour des déficits mnémotiques ou cognitifs. Cette thèse cherche donc à combler ces lacunes. Grâce à des études de cas sériés, les déficits linguistiques de 11 patients APP sont analysés dans un cadre neuropsychologique cognitif et comparés aux déficits linguistiques d'un groupe de 11 patients DTA. Plusieurs instruments ont été élaborés ou peaufinés pour pouvoir tenir compte des détails de l'analyse, notamment un système de production conçu pour identifier les dissociations fonctionelles du système linguistique d'après une analyse de la performance des patients à partir d'une version informatisée de la Psycholinguistic Assessment Battery (Caplan & Bub, 1990). Même si le résultats donnent à penser qu'il existe une grande hétérogénéité des symptômes au sein de la populations APP, plusieurs déficits susceptibles d'être utiles pour poser un diagnostic différential sont examinés de plus près, notamment les déficits de lecture, la compréhension des termes abstraits, la compréhension auditive, le traitement des affixes et l'accès sémantique. Enfin, l'importance des résultats pour les modèles actuels du fonctionnement linguistique fait l'objet d'une analyse plus approfondie.

Table Of Contents

Acknowledgments

During the course of this research, over seven thousand (!) data files were generated, many of which needed to be moved, renamed, and/or re-formatted multiple times. I would never have been able to manage the data without Userland's scripting application, Frontier, and CE Software's macro software QuicKeys. This is not a paid advertisement.

Similarly, this project would have been finished before it began if Dr. Peter Shell, at CMU, had not made his Common Lisp frame system, Parmenides, and his production system shell, FRuleKit, available free of charge. His generousity in making that code available, and his patience in answering my many naive questions about how to use it, are greatly appreciated.

I'd like the thank the many subscribers to the 'info-mcl' mailing list who generously gave time to a stranger to help me solve a number of programming problems that seemed insoluble.

Ingrid Johnsrude, Diane Kampen, and David Sinyor all read early versions of this thesis. I appreciate all their comments. Other readers of this thesis may also want to thank them, for convincing me to move the case studies from the main text to two appendices.

Rhonda Amsel always gave me 'the right answers' to statistical questions, which I appreciate.

Geoff Schultz saved me from insanity on more than one occasion by proving that statistics do indeed work in a manner which is consistent with the laws of probability, and generously took time from writing his own thesis to write me a program for calculating Fisher's Exact Probability. I also benefited from many conversations with him about this thesis, and appreciate his listening to me.

Special thanks are due to all the neurologists who provided me with patients, and to David Caplan for providing me with normal data collected at his lab. I'd particularly like to thank Howie Chertkow, who gave me much additional help and support.

Genevieve Gore and Phillip Jackson helped collect some of the data as part of their undergraduate studies. They bore with me through a time of considerable confusion and technical difficulty, and asked me pointed questions which clarified my thinking. They both have my thanks for doing so.

Irv Binik helped me understand what psychology was all about, and was the catalyst for my continuing interest in both diagnostic issues and expert systems. I appreciate his help and support over many years.

Special thanks to Mathew Shapiro for stepping into the breach.

The patients in this study made it possible. I greatly appreciate their patience and effort, as well as the opportunity they provided for me to learn so many important lessons about courage and grace under difficult circumstances. My thanks go out to all of them.

Dan Bub has been an unfailingly gracious, generous, and good-humoured thesis supervisor, putting up with me unflinchingly even at my darkest and crankiest moments, and going far beyond the call of institutional duty on innumerable occasions. He has been a mentor in the best sense of the word, having taught me far more than just cognitive neuropsychology. I'm the better person in many ways for having worked with him. Thanks, Dan.

My wonderful friend and wife, Elena Nicoladis, deserves more thanks than I can fit into a work of less than encyclopedic size. She bore the brunt of the many mutant forms of angst which evolved in my psyche during the course of my doctoral studies, never wavering in treating them all with the gentle and loving indifference they deserved. Thank you, Elena.

Along with Dan and Elena (and too many others to mention here) I'd like to thank Spalding Gray, Michel De Montaigne, Henry Miller, Tom Robbins, and Ludwig Wittgenstein for their eloquent clues about how to maintain that peculiar perspective on life without which no work can be worthwhile.

Finally, I'd like to thank my parents, Robert and Clare Westbury, who have always supported the circuitous, slow, and bizarre route I took through academics to get to the stage of writing a doctoral dissertation. They must often have wondered if any concrete fruits would ever be produced. Their easy acceptance of the strange life style I have chosen and their life-long encouragement (or at least tolerance!) of my often inane intellectual curiousity made the process easier for me. I appreciate this very much. Thank you.

If there is any merit to this work, I dedicate it with love and gratitude to the two of them.

Prologue

One of my favourite parables is Chuang Tzu's oft-repeated tale of Prince Wen Hui's encounter with his Taoist cook. After he has been admired by his employer for his skill at cutting up an ox, the cook declares that his apparent skill is simply a reflection of his ability at finding the natural cleaves in the ox's flesh. In Thomas Merton's (1965) poetic translation of the tale, he explains:

> There are spaces in the joints;
> The blade is thin and keen:
> When this thinness
> Finds that space
> There is all the room you need!
> It goes like a breeze!...
> True, there are sometimes
> Tough joints. I feel them coming,
> I slow down, I watch closely,
> Hold back, barely move the blade,
> And whump! the part falls away
> Landing like a clod of earth.

I have tried to take an approach to dismembering the human language system that is very similar to the approach that the Taoist cook took towards his own work. I too have sought to be 'guided by the natural line' of the whole, to find the natural lines along which that whole could be cleaved. Like him, I have encountered some tough joints which have slowed me down (but which have perhaps not, alas, always yielded to my patience!). Like him:

> After three years,
> I no longer saw this mass.
> I saw the distinctions.

According to Chuang Tzu's tale, it had taken the cook nineteen years to perfect his apparently simple skill. I have been trying to perfect my own skill as a cognitive psychologist for many years fewer, and cannot yet lay claim to any such mastery. Nevertheless, today I do allow myself to share briefly in something of the contentment of the prince's cook:

> I withdraw the blade,
> I stand still
> And let the joy of the work
> Sink in.
> I clean the blade
> And put it away.

July, 1995

Chapter One: Introduction

Every living being is also a fossil.

Jacques Monod
Chance & Necessity

Thirteen years ago, Mesulam (1982) published a report describing six patients manifesting a slowly progressive aphasia without any accompanying signs of dementia. Although such cases had been previously reported in the literature (Poeck & Luzzatti, 1988), it was the publication of Mesulam's cases which led to the recognition of a syndrome characterized by pure aphasia without dementia, now known either as Mesulam's Syndrome or (more commonly) Primary Progressive Aphasia (PPA). The existence of PPA has been widely documented in the years since Mesulam's initial publication. Unfortunately, the existing published literature is unsystematic, consisting largely of single case reports which have relied upon a bewildering variety of different neuropsychological instruments, many of which were developed for other purposes (Caplan, 1992). These limitations have made it difficult to put the published results into a coherent theoretical structure. It is perhaps due to this lack of coherence that almost no work has been done which explicitly attempts to discover if the nature of the language deficits in PPA allow it to be easily distinguished from the most common dementing disorder, and the most important differential diagnosis, dementia of Alzheimer's type (DAT). The distinction between DAT and PPA was never made before 1981. Even today only a few medical practitioners who have a particular interest in language disorders routinely differentiate between the two disorders. Nevertheless, the distinction is an important one. DAT has a rather different prognosis as well as a different neural mechanism. Like PPA, DAT almost invariably affects language function, as we shall see in Chapter 4. Since language problems can masquerade, to both the patient and the psychologically-naive observer, as memory problems, a proper understanding of the role of language deficits in both PPA and DAT is vital to making the diagnostic distinction. This thesis is an attempt to add to this understanding by examining the nature of language deficits in both disorders in systematic detail.

Along with its diagnostic relevance, there is another good practical reason for the intense interest that PPA has aroused in cognitive scientists interested in understanding language, interest that might seem at first glance disproportionate to the frequency of PPA or to the importance of distinguishing two disorders which currently can be neither controlled nor cured. The reason for the interest is that PPA has the potential of providing a unique source of clean 'natural experiments', which can offer us insight into how the human language system is functionally organized. Unlike DAT patients, those patients with pure PPA do not have extra-linguistic cognitive deficits which are likely to impact on their performance on language tests. In this respect, PPA patients do not differ from some patients who have suffered language disorders secondary to strokes, who have been the main source of natural experiments to date. Where they may well differ is in the range and type of experiments allowed by the source of the damage. Strokes do not occur randomly throughout the brain, but tend to cluster around areas where an artery either turns sharply, or branches, especially when the off-branch is much smaller or thinner than its parent vessel (Mohr & Kase, 1983). Since occlusion of a branch of the middle cerebral artery (which supplies blood to both Wernicke's area and Broca's area) is the leading cause of aphasia (Barr & Kiernan, 1983; Poeck, 1983), it is not surprising to find that most aphasic patients who have been studied in the published literature acquired their aphasia following a stroke. Although the literature realistically reflects the population base rates for acquired aphasias, this bias towards stroke-induced aphasias has to date presented a wholly practical limitation on the knowledge we have been able to gain from patient studies about how language is actually organized in the brain. The theoretical inadequacy of most clinical aphasia batteries is in part due to the fact that such batteries have usually been designed to measure those only those decompositions of the language system which one might expect from stroke damage, rather than being organized along more general psycholinguistic principles (Caplan, 1992).

PREVIEW

The three sections of this chapter introduce the problems this work will address. In the first section, the problems of aphasiological research are placed in their historical context, since, without some understanding of this context, it may be difficult to understand the methodological approach that has been adopted. In the second section the hypothesis which were formulated when this work was begun are presented and justified. In the third section, some of the methodological contributions this research makes are introduced, by describing the tools which were developed in the course of the research described in this thesis. In Chapter 2, both those tools and the methodology of this study are described in greater detail. Chapter 3 reviews previous studies of PPA. In Chapter 4, the PPA data collected for this study (which is presented, along with case descriptions, in Appendix B) are analyzed and discussed. In the first section of Chapter 5, a brief review of previous studies of linguistic deficits in DAT is presented. In the second section, the DAT data collected for this study are analyzed and discussed in the same framework as was used in Chapter 4. The DAT case descriptions are presented in Appendix C. In Chapter 6 the linguistic deficits of the two disorders are compared and contrasted. In the final chapter, the results are summarized, and future directions for the work are presented.

HISTORICAL CONTEXT

Language, Science, and Cognitive Neuropsychology

The study of language and its relation to the architecture of the brain has long presented particular difficulties for scientists. The cognitive neuropsychological approach used in this thesis is the end result of a century-long search for a practical solution for these difficulties. In this section, the approach is placed in context is order to clarify both the nature of the questions this work addresses, and the nature of the answers that can be reasonably expected to be found.

There are a great many reasons why the neuropsychological study of language is peculiar compared to the study of the other widely-studied neuropsychological functions, such as attention, memory, and visual processing. Among others, language researchers face the following problems:

i.) They must do without animal lesion studies, relying instead upon natural lesions which are often large and imprecise;

ii.) Since language is only defined at a high level of abstraction, they must forego the 'low level' analyses which have been fruitful in other subfields;

iii.) They must try to disentangle massive functional interdependencies which are largely controlled by a relatively small area of cortex and which often cannot be easily broken up into meaningful subfunctions;

iv.) They must deal simultaneously with multiple interconnected input systems, and with output which is infinitely flexible;

v.) They must control for a number of known relevant variables which now numbers in the dozens;

vi.) They must accept the fact that subjective 'meaning' (semantics) rather than purely deterministic mechanisms in receptive tissue account for a great deal of the variation with which that they are faced.

Over the past one hundred years, the attempt to deal with problems such as these has led to the development of a unique approach for studying the functional architecture of language. Since it is not possible to understand this approach without understanding something of its long and complex history, I begin here with a brief summary of that history.

If we disregard models proposed prior to Broca's localization work as pre-scientific (a dubiously chronocentric but widely accepted claim) then the earliest serious attempt to model the structure of language system was the model proposed in 1874 by Wernicke. As might be expected from a pioneering and early effort, Wernicke's model was not based on the controlled testing of any theory, but rather developed as an attempt to systematically explain a seemingly disparate set of clinical observations. His theory was modeled as a diagram of connected modules, organized in such a way that lesions between the modules in the model could explain clinically-observed sets of deficits. Though some features of Wernicke's model were criticized and rejected by some aphasiologists (including a young neurologist named Sigmund Freud) his model was generally hailed as a major achievement. Wernicke's approach to aphasiology became accepted as a mainstream scientific paradigm.

However, in the decades which followed, the 'diagram makers' (as they were referred to with apparently sarcastic intent by an early critic of the approach, Henry Head) fell out of favour. There were two main reasons for their fall. The first main reasons was that their implicit claim that the functional components of the models had precise analogs in the brain did not stand up to scrutiny. Based as they were upon crude and idealized psychological concepts, the kinds of pure deficits which their models predicted were never seen. The second main reason for the fall of the diagram makers was that the only evidence they could bring forward to buttress their claims was often flimsy, cursory, and anecdotal. Their claims were therefore difficult to replicate and, sometimes, difficult to accept as objective. Thus Head was able to claim that the diagram makers would "lop and twist their cases to fit the Procrustean bed of their hypothetical conceptions" (cited in Marshall's introduction to Howard and Franklin, 1988, pp. ix).

The scientific failure and subsequent rejection of the diagram makers was not accompanied by a rejection of the diagnostic categories they had created. The aphasic subtypes that they had proposed endured because those subtypes were viewed (perhaps unfortunately) as clinically-useful heuristic devices even when they no longer had any theoretical value. Because the categories were no longer tied to a specific theory, there was no principled reason to change them even when they were found to be imprecise. New characteristics which were observed clinically to be sometimes associated with one of the classical diagnostic categories were simply noted as possible but not necessary features of that category. As Schwartz (1984, pp. 6) has noted, the "resulting list [of aphasic subtypes] has a decidedly eclectic, ungainly look to it; anatomical features are considered alongside circumscribed linguistic deficits and more diffuse behavioural abnormalities." Due to these limitations, it became very difficult to formulate any systematic understanding of what was happening when a person became aphasic.

Despite its initial failure, diagram making did not die. For many years it played a much less important role in the field, which as a result "became much less interesting" (Shallice, 1988, pp. 13). The revival of diagram making in cognitive neuroscience in the last three decades may be attributed largely to the rise of the information processing paradigm, whose rise was in turn due to a complex set of technological and theoretical advances in a number of related fields (see Gardner, 1985, Chapter 5). The information processing paradigm was naturally suited for modeling with diagrams. However, the new diagram makers had learned from history. They placed a renewed emphasis on the case study. Those studies were no longer supported with the subjective clinical reports which had lead in part to the downfall of their predecessors. They were rather backed up with rigorous testing procedures which allowed for objective description and which were also carefully controlled for known factors (e.g. word characteristics such as frequency) which had been shown in other work to affect word processing abilities. The new breed of diagram makers had also learned from their predecessors not to engage in endless taxonomic disputes.

By the mid-1980's, there was widespread agreement within the field that classification of aphasics by type should be done away with altogether (Ellis, 1987; Badecker & Caramazza, 1986; Coltheart, 1985; Schwartz, 1984; Caramazza, 1984). A corollary of this rejection of group labels was a rejection of group studies, with a resultant increased emphasis on the case study. This emphasis was not intended to imply that researchers had to limit their study necessarily to a single patient, but rather to stress that the unit of analysis, even in studies with many subjects, must be not group but individual performance. Of course, group studies continued to be recognized as necessary for establishing normal performance, and as useful in other situations in which there are principled reasons for assuming neuropsychological equivalence within a group.

To be taken seriously as a scientific discipline, a research methodology must offer a way to generalize from the specific case. Cognitive neuropsychology does in fact generalize. However, its generalizations are made across groups for which membership criteria is defined in a primarily theory-driven rather than a primarily data-driven manner. The theory which specifies to which group an individual patient belongs usually requires an extensive and detailed analysis of his or her individual pattern of responses.

There are many reasons why such a fine-grained analysis is considered necessary. I will not present a detailed discussion of these reasons here. However, a simple example can make the general nature of the problem clear.

Consider the performance of patients X and Y on two different tests A and B, as illustrated in Table 1.1. Let us assume that the normal average score on both tests is near ceiling, with a very low standard deviation. When the two patients are considered together, they obtain a low and equal average score (55%) on both tests. However, it is clear from inspection that the two patients performed very differently. Patient X was very poor at Test A, but scored in the normal range on test B, while patient Y shows the inverse performance. If we have reason to believe that performance on test A is independent of performance on test B- that is, if the two tests results are doubly dissociable- then patient X should not be considered in a group with Patient Y, but rather in a group with other patients who show a similar pattern of performance on the entire set of relevant tests. We can only know if we are dealing with a patient like Patient X or one like Patient Y after we have given our patient the relevant tests. It is only when we have ascertained what kind of patient we are dealing with that we may make statements about group performance and relative frequency for that type of patient, statements of the form "100% of patients who were like Patient X also scored at least two standard deviations below norms on another test C" or "27% of patients from a given diagnostic group showed the same pattern of deficits as Patient Y." Many more interesting and subtle examples of similar kinds of problems that arise from masking the role of individual variables are discussed in Shallice's (1988) book, *From Neuropsychology To Mental Structure.*

	Test A Score	Test B Score
Patient X	10%	100%
Patient Y	100%	10%
Average	55%	55%

Table 1.1: Hypothetical Data
Of Two Patients On Two Tests

The rejection of the early taxonomy and the resultant rise of the case study was derived from a growing appreciation of the difficulty of defining syndromes in terms of information about the statistical coherence of symptoms. That approach had been shown to be a failure, leading as it did to one of three equally undesirable ends: to the messy eclecticism derided above by Schwartz (1984), to a dissolution or fractionation of syndromes once thought to be stable, or to an outright oversimplification of the facts. It became clear that any such taxonomy was destined to fail unless it could be founded upon a full and explicit understanding (at a functional if not a neurological level) of the causal mechanisms underlying the symptoms that define the syndrome. The attainment of such understanding remains the guiding aim of the field.

No full understanding of the causal mechanisms underlying a language deficit is yet possible, since we have only a weak understanding of the neural mechanisms which underlie normal cognitive processes, and no understanding at all of how (or if) lesions in the brain may modulate the very parameters which govern cognitive performance in normal subjects, rendering patient performance qualitatively different from normal performance. Cognitive neuropsychology today is in the same situation as was zoology prior to Darwin (or, more precisely, prior to the rediscovery of Mendelian genetics in the 1860s): we may believe that there is some order to the bewildering variety of cases we are faced with, and we may have faith that this order is founded in a deterministic organic mechanism, but as yet we have absolutely no idea how to go about teasing that deterministic mechanism out. Like the early biological taxonomists, we do not yet know what kinds of variety might be most directly coded by the underlying controlling mechanisms. We are waiting for contributions equivalent to those made by Gregor Mendel to biology: for concrete and indisputable evidence of what form the order which is sought may take. Until we have such evidence, all variety must therefore be looked upon as equally important, even though we know that doing so leaves us open to dizzying growth in the number of possibly relevant interdependent characters we will want to track in our attempts to understand.

With this historical overview, the goal of the present work may be more easily understood. The present work was undertaken with two main goals in mind. The first goal was to present as detailed a description as possible of the symptoms of a particular language disorder, PPA. In doing so, I was not eschewing the opinion of the group cited above, that the classification of aphasics by type should be done away with altogether. Since a diagnosis of PPA is not dependent upon the type of language deficit a patient has, but rather on the relation of the language deficits to other symptoms, and on the variation of those deficits over time, the diagnosis may be made without classifying a patient's type of aphasia. In this thesis I have not made any attempt to classify the patients according to any extant taxonomy of aphasia. I have rather tried simply to describe their language deficits in as much detail as possible, with as many controls as possible, in an attempt to understand in as precise detail as possible what is wrong with the language systems of the patients in this study. The nature of the analysis presented here is motivated by known (or, in a few cases, suspected) functional fractionations of the language system. Information about the ways the language system can fractionate is explicitly built in to the structure of the PAL tests, as will be seen in the next chapter.

The second goal of this work is to try to exploit the strength of the case study approach, which provides another way than the classic hypothetico-deductive method (see Popper, 1959, pp. 40-44) of making contribution to scientific knowledge: the method of existence proof. Carl Jung wrote that "One does not need to produce ten thousand duckbilled platypi in order to prove they exist" (Collected Works, 18:1198, cited in Wilmer, 1987, pp. 171). Similarly, one does not need to produce ten thousand patients with a specific set of functional impairments in order to prove they exist. A single aphasic patient provides indisputable evidence about one possible way the language system can break down. As the number of patients with a similar set of symptoms increases, it becomes increasingly likely that a set of symptoms reflects something systematic about the way the language system is arranged, rather than simply reflecting the effects of an arbitrary set of independent lesions.

Although the method of existence proof may be considered suspect in science because the only explanations it can offer are post-hoc, it can play a vital role in science in two situations. The first situation is when theories are stated (as Popper's 1959 prescription for science insists they should be) in explicit and clearly falsifiable terms. When a theory is well-stated, then a single observation which violates the theory can disprove it. The insistence by practitioners upon reliance on single case studies should be (but is not always) accompanied by an equal insistence upon statements of hypotheses which allow for a clear and simple refutation (see Caramazza, 1986). We shall see an example of such an hypothesis later in this thesis, when we offer evidence which refutes one. Although the present work does not formulate a theory about how language is instantiated in the brain, or refute any existent theory in detail, it will shed some light on the kinds of associations which are likely among deficits, and on the question of what kinds of deficits a theory of language must be able to explain.

The second situation in which the existence proof method of science can play a vital role is when the correct unit of analysis is not clear, and so the statement of a clearly falsifiable hypothesis is not possible. This was the situation of zoology prior to the last century, and chemistry in the century prior to that. It is also the situation of cognitive neuropsychology today. I call the problem presented by this situation as 'the problem of the hidden variable', to emphasize that the problem is that the units of analysis which are apparent may not be, and in many cases are actually unlikely to be, the relevant units over which a falsifiable theory can be defined. Until one truly knows how to define the class of items about which one wishes to make an hypothesis, one cannot tap the power of the hypothetico-deductive model of doing science.

Although there are many disputes of detail in the field, the general logic underlying the method of existence proof as it relates to functional dissociations and associations is well worked out (see, for example, Shallice, 1988; Caramazza, 1984, 1986).

This thesis also attempts to clarify the units over which a theory of linguistic processing must be defined. As will be seen in later chapters, the data collected for this study are especially useful for providing a preliminary and circumscribed examination of two aspects of language which have been largely neglected in current models of language processing: the way that processing of different word types (e.g. different frequencies, orthographic and phonological complexities, or syntactic roles) dissociates in language system, and the role of word affixation (the creation of new words by adding affixes to root forms of a word). Not only has this latter aspect of language been relatively little studied, but (as pointed out in Caplan, 1992) there is virtually no work to date which examines how processing of compound word forms is affected by neurological disease.

HYPOTHESES

Before I began this work, I formulated a number of hypothesis, which are reviewed in this section.

Hypothesis 1

The main hypothesis is a simple one, which may be stated in either a strong or a weak form:

1a.) Strong version: There is a systematic difference in terms of primitive psycholinguistic operations between patients with PPA and patients with DAT. This difference will allow one to decide with certainty which disease a person has given that patient's performance on tests which focus on the relevant primitives.

1b.) Weak Version: There are probabilistic differences in terms of primitive psycholinguistic operations between patients with PPA and patients with DAT. These differences will allow one to decide with an empirically-grounded degree of probability which disease a person has given that patient's performance on tests which focus on the relevant primitives.

This hypothesis is defined in functional terms, but founded upon the neuropathological differences which have been noted between DAT and PPA patients. The remaining hypotheses are founded upon a more detailed consideration of the documented differences in brain pathology between patients in the two groups. Such a foundation is the only one which was possible prior to the completion of this work. There have not been enough detailed neuropsychological studies of PPA patients to allow for the formulation of defensible hypotheses formulated in wholly functional terms. Such hypotheses are, for reasons which will become apparent in this thesis, more desirable.

It is very important to understand that the decision to base the following hypotheses upon crude neurological dissociations should not be seen as an endorsement of any theory of precise localization of function. After all, a recognition of the problematical nature of such theories (well documented in Caplan, 1992; Shallice, 1988; and Steinmetz and Seitz, 1991) is one of the defining forces of the cognitive neuropsychological tradition in which this thesis situates itself. The hypotheses which follow are all formulated in a particular manner, using information about neuropathological differences between DAT and PPA as a rough guide to buttress hypotheses which are stated and testable the level of functional decomposition. The decision to formulate hypotheses in this manner was made after long consideration of a problem that was first elucidated by the cybernetician Gregory Bateson (Bateson, 1979. The interested reader should also consider Bateson, 1972; Bateson, 1991; Watzlawick, 1984). Bateson had particular derision for hypotheses which he called 'dormitive principles', in honor of the doctor in a play by Molière who claimed that opiates caused sleep because they contain a dormitive principle. Dormitive principles are assertions which introduce a circularity in the same way that the physician's explanation did, by assuming in an explanation the existence of the very entity one is trying to explain. Dormitive principles are extremely common, an almost inevitable result of the widespread belief that science consists merely of fact-collecting, hypothesis formulation, prediction from hypothesis, and examination of how the hypothesis accords with the facts. This approach to science implicitly instantiates the dormitive principle, by failing to specify anything about the relation between the facts of observation and the facts of hypothesizing. Bateson (1991, pp. 171) has stated (and I agree) that "about three-quarters of all the hypotheses in the behavioural sciences are fundamentally dormitive principles".

Dormitive principles are appealing for many reasons, not least of which is that they are compatible with the real pressures (both political and psychological) which have forced scientists to put a high value on prediction. After all, it is usually only when we predict correctly that we have a publishable result. But, as Bateson (1972) notes, prediction is a poor- or at least, an uninteresting- test of a hypothesis when the hypothesis is a dormitive principle, for the simple reason that such an hypothesis has been formulated precisely in order to be predictable. One can multiply dormitive principles by fractionation for a very long time (for an entire, career, easily) fulfilling predictions all the while, without ever having escaped from the fundamental circularity of the whole enterprise. The rich histories of both behaviourist and psychoanalytic theory (both founded on a complex set of dormitive principles) attest to this. This process of multiplication of dormitive principles, says Bateson (1972, pp. 26) "must always lead to something like the present state of the behavioural sciences - a mass of quasi-theoretical speculation unconnected with any core of fundamental knowledge".

It is easier to state the problem of the dormitive principle than it is to state the solution. In his writing, Bateson suggests that the way to avoid dormitive principles is to divorce the fact collection process from the hypothesizing process. This does not mean that scientists must hypothesize in a vacuum. What it does mean is that a hypothesis must be limited not by facts at the same logical level as the facts which will be collected as data, but rather by a tentative theory (a set of propositions which are either axiomatic or tautological) guided by principles which are stated at a different logical level than the level at which data will be collected. The difficulties inherent in using limitations at one level to guide theorizing at another level forces one to be very careful in stating such principles.

The easiest hypotheses to state in taxonomic work are strong dormitive principles. For example, any hypothesis which infers from the observation of a set of prior patient descriptions to one's own patient population is a dormitive principle. Such hypotheses are hypothesizing about *description*, not about explanation- and thus they are not really hypotheses at all. The whole point of an hypothesis is to test an explanation. Hypotheses about description do not (and can not) do so. They are merely imitative: scientistic rather than scientific; comforting if one wants to feel one is acting like a scientist, but ultimately useless. Insofar as my aim is merely to describe what I see, I do not need any hypothesis at all. I need merely know how to formulate a description.

My aim in formulating the following hypotheses was to attempt to escape from dormitive principles by placing the hypotheses stated below in the context of known data about lesion sites in PPA and (especially) DAT, and how those lesion sites are related to deficits. I focused on the lesions associated with DAT rather than the lesions associated with PPA because the affected areas in DAT are more distinct, better documented, and more easily functionally-dissociable than the more circumscribed affected areas in PPA. In formulating these hypotheses, I was well aware that the current state of knowledge in relating functional deficits (especially of language use) to the underlying neuropsychological substrate remains crude. Many of the alleged relations upon which the following hypotheses depend are therefore themselves hypothetical.[1] I believe that others will eventually be proven to be over-generalizations of relations which will then need to be specified in much finer anatomical and/or neuropsychological detail. For all of these reasons, my only purpose in formulating these *a priori* hypotheses was that they provide some focus for the ensuing investigation by giving it some reasonable starting points. In later chapters, the data presented will simultaneously force and allow consideration of relations which go beyond these crude hypotheses.

[1]As an aside, I am convinced (following Quine, 1961) that this 'mutual testing' relationship always holds true between a hypothesis, and the theory which motivates that hypothesis. I also believe that the relationship is under-appreciated in science. If a hypothesis which is well-formulated (in Bateson's sense, as described above) fails to find support, one cannot always know if the failure came about because the theory upon which the hypothesis was based was incorrect- or if it came about because, as we usually assume, the hypothesis itself was incorrect. There is a subtle question here, regarding where one makes ones epistemological stand and why - or, equivalently, a question regarding which of two sets of variables shall be the independent and which the dependent variables. (Thanks to Dan Bub for making me aware of this equivalence.) I will not discuss this intriguing problem further here.

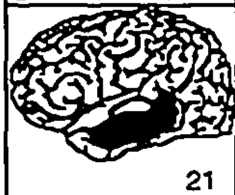
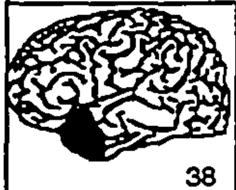
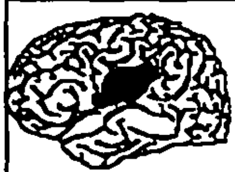
Brodmann's Area	Location	% Loss	Associated Deficits
 21	Middle temporal gyrus	46%	Phoneme discrimination; naming; verbal comprehension; repetition; reading irregular words ('reading by sound');
 38	Anterior tip of temporal lobe	48%	None recognized
 10, 46 (grey)	Pre-frontal	42%	Semantic search / prelinguistic concept formation
 40	Supra-marginal gyrus / Caudal half of first temporal convolution	50%	Conduction (Wernicke's) aphasia: speech production errors; phoneme discrimination; reading nonwords ('reading by sight vocabulary')

Table 1.1
Language-region areas of cortical cell loss in DAT, and their associated functions (after Bayles, 1987, pp. 9, McCarthy & Warrington, 1990, Steinmetz & Seitz, 1991)

The main sites of relevant cortical degeneration in DAT patients, and the performance deficits associated with damage to those areas, are summarized in Table 1.1. The analogous sites in PPA patients are examined in some detail in Chapter 3. For now, we may content ourselves with the following brief summary: degeneration in PPA patients is almost always seen in the left temporal lobe, most commonly and clearly in the perisylvian region; it seems to be limited to the posterior portion of the frontal lobe and the anterior portion of the parietal lobe; and it is (with a single known exception) never seen in the parietal lobe unless it is also in the temporal lobe. We can infer from all these facts that the degenerative process associated with PPA probably begins in either the posterior frontal or the temporal lobe, and spreads posterior to the parietal lobe along the Sylvian fissure: that is, along the superior temporal gyrus.

Hypothesis 2

Based on this predominance of degeneration in PPA patients mainly along the Sylvian fissure, one can expect a strong overlap between the two disorders in deficits associated with area 40, a site where degeneration has been found in both disorders, while finding a dissociation of deficits associated with area 21 (more inferior in the temporal lobe), which is strongly associated with degeneration in DAT but not in PPA. Thus I hypothesized that:

> 2.) DAT patients will show greater deficits than PPA patients in tasks known to be modulated through area 21: accessing names in the picture-naming tasks; verbal comprehension; repetition; reading irregular words; and writing. Both DAT and PPA patients will show deficits on tasks known to be modulated through area 40: speech production; phoneme discrimination; and reading nonwords.

Hypothesis 3

Degeneration in PPA often extends to the frontal lobe. However, the main site of frontal lobe degeneration in DAT patients is in Brodmann's areas 9, 10 and 11, at the extreme anterior of the lobe. Lesions to this area are not generally recognized to have any effect on language per se, although they are implicated in concept formation (Steinmetz & Seitz, 1991). However, there are of course important language areas in the posterior frontal lobe, most notably Brodmann's areas 44, 45, and 47, collectively referred to as Broca's area. Damage in this area has long been known to cause deficits in the motor production of language, and in comprehension of sentences, especially when that comprehension depends upon the utilization of syntactic structure, rather than on lexical or pragmatic considerations (McCarthy and Warrington, 1990). Because damage here appears more likely to be due to the PPA degenerative process than to the DAT process, I hypothesized that:

> 3.) PPA patients will show greater deficits than DAT patients in tasks known to be modulated through Broca's area: tasks involving oral production of language, and sentence comprehension, especially of semantically reversible sentences.

Hypothesis 4

I also hypothesized that there would be differences between patients in the two groups in the nature of their graphemic output deficits. Patients with lexical agraphias (who rely on common sound-to-spelling correspondences and thus cannot spell irregular words) have been found to have lesions in the parietal lobe, probably at the junction of the posterior angular gyrus, and the parietal-occipital border (Warrington and McCarthy, 1990). We would expect patients with the appropriate lesions to make errors in writing irregular words on the written naming and written sentence production tests. The parietal lobe is not an area associated with major neuronal loss in DAT, but degeneration in PPA is known to often extend posterior into the parietal lobe. Therefore it seemed reasonable to hypothesize the following:

> 4.) Because they are more likely than DAT patients to have neuronal degeneration extending into the parietal lobe to the junction of the posterior angular gyrus, and the parietal-occipital border, PPA patients will make more orthographic errors in tests of writing than DAT patients.

Hypothesis 5

A somewhat more complex hypothesis was made regarding expected differences in visual recognition. Visual recognition is not a unitary neuropsychological module, but relies upon intact functioning of several modules. To name an animal, a person must be able to parse visual form (that is, to correctly pick out the visual features of a stimuli), to relate the form to other recalled forms, to relate those recalled forms to prior experience and wider knowledge, to access a lexical label for the form, and to map that label onto its phonemic or orthographic form for motor (spoken or written) output. These functions are known to be largely dissociable (see Chapter 12 of Shallice, 1988). We might think of the naming process roughly as a mapping into three quasi-independent spaces (a semantic space, a lexical space, and a phoneme or grapheme space), as illustrated in Figure 1.1.

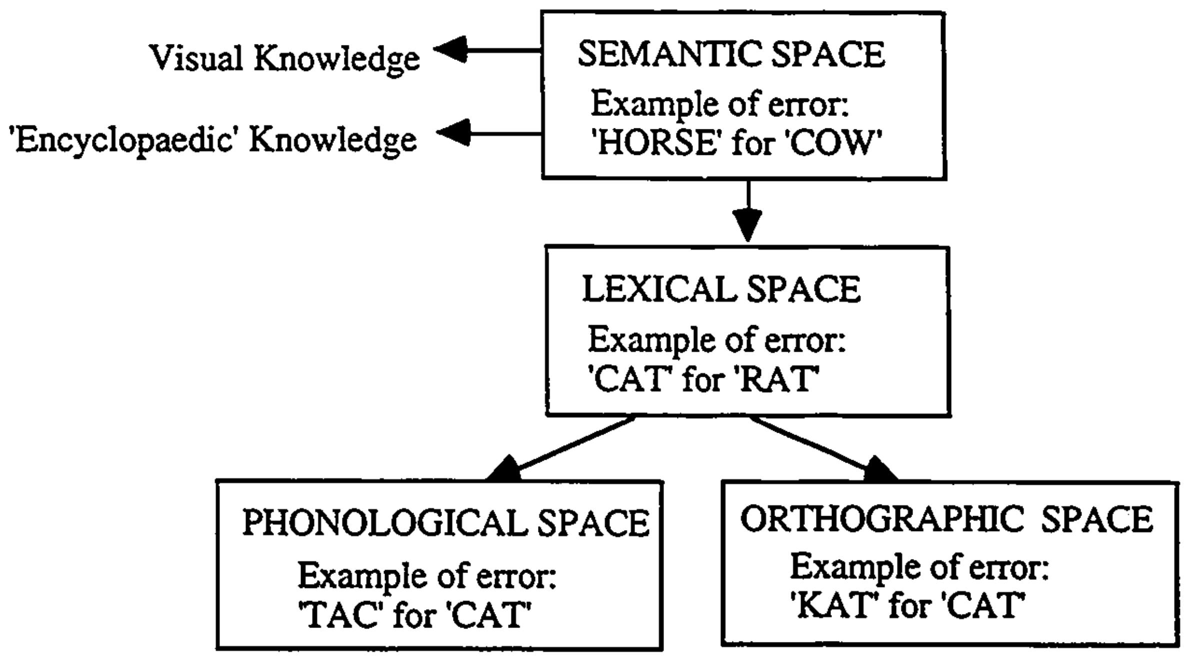


Figure 1.1
The spaces which must be mapped into in order to name a picture.

Damage to pre-frontal cortex (Brodmann's areas 9-11, 46, and 47) has been associated (Steinmetz & Seitz, 1991) with a well-documented deficit in DAT patients (Irigaray, 1967; Warrington, 1975; Constantinidis et al, 1978; Schwartz et al, 1979; Obler, 1981; Appel et al, 1982; Martin & Fedio, 1983; Flicker et al, 1987): an impairment in searching through semantic space, or an impairment in 'associative fluency'. Patients with pre-frontal degeneration tend to make 'true' naming errors: that is, naming errors which are semantically mediated rather than being secondary to other impairments in language production. Brodmann's area 10 is one site associated by Bayles (1987) with major degeneration in DAT. In PPA patients, however, the degeneration does not tend to extend frontally, thus missing one major area associated with the semantic deficits. We therefore do not expect to see deficits at the level of relating perception to stored semantic knowledge. However, as noted above, we can expect to see deficits in PPA patients in production of names, especially spoken production, without the associated loss of semantic knowledge. Although the complex nature and underlying neuroanatomy of visual object recognition renders hypothesizing about it extremely difficult, as an orienting guide it seemed reasonable to hypothesize that:

> 5.) DAT patients will tend to be more impaired than PPA patients on the semantic battery tests which use both written and picture stimuli, because these tests rely upon access to stored information about objects. PPA patients will show difficulties in naming pictures (especially in the oral modality) with normal performance in accessing associated information, as measured by the semantics battery.

METHODOLOGICAL CONTRIBUTIONS

Useful scientific work may often be achieved by bringing new tools to bear on the phenomenon of interest. Consider, for example, how fruitful were the inventions of the microscope (Boorstin, 1983), Watson & Crick's templates for representing the structure of DNA (Judson, 1979), or even, to choose a slightly more controversial example first suggested by the philosopher Ludwig Wittgenstein (Monk, 1990), the terminology of classical psychoanalytic theory. These very different inventions did not 'merely' allow questions to be answered which had previously been impossible to answer. Each one of them also allowed questions to be posed which had previously been impossible to pose. As the poet e.e. cummings (1961) once wrote "Always the beautiful answer who asks a more beautiful question." The usefulness of new ways of gathering or understanding data is particularly strong during the early stages of a field's growth, when the relevant terminology, concepts, and representations are not yet developed enough to allow for the appropriate formulation of scientific hypotheses.

Because of the complexity of the relationships between the data gathered for this thesis, the work described here was made practically possible only by the development or refinement of four important tools. The four are:

i.) A computerized version of Caplan & Bub's (1990) Psycholinguistic Assessment Of Language (PAL) Battery

ii.) A production system for analyzing the results from that battery, along with:

iii.) A secondary knowledge engineering tool, Janus, which was created in order to make the development of that analysis system practically feasible

iv.) A new notational system which was developed to represent the results from the production system in a way that rendered those results comprehensible.

The first of these tools, a computerized version of Caplan & Bub's (1990) PAL battery, was refined and restructured quite extensively for both practical and theoretical reasons. I also calculated the normal statistics (using data collected in Boston by Caplan, but never previously analyzed) and an item analysis for each of the tests. My refinements and comments, as well as constructive criticisms which are derived from analyses of the normal data, are presented in the next chapter. The remaining three tools have not been previously described in the literature. They are therefore also described in Chapter 2.

Chapter Two: Methodology

As God's truth is what God comes to know as he creates and assembles it, so human truth is what man comes to know as he builds it, shaping it by his actions. Therefore science (scientia) is the knowledge (cognitio) of origins, of the ways and the manner how things are made.

Giambattista Vico
De Anquissimia Italorum Sapientia

In this chapter, the general methodology of my study is reviewed in detail. In the first section, I describe and evaluate the main tools (including the tests upon which the study depended) which were developed specifically for the work. In the second section, I explain the actual methodology of the study. In the final section, I address some potential criticisms of that methodology.

TOOLS

i.) The PAL Battery

The original PAL Battery (Caplan & Bub, 1990; Caplan, 1992) consisted of 27 subtests. The number of distinct subtests is somewhat misleading from the point of view of conceptual understanding of the battery, since several tests were designed in such a way as to allow for the testing of more than one single function which was known to be dissociable, and many tests contain information which can only be interpreted in the light of performance on other tests. The rationale behind the choice, design, and original organization of the tests is explained in some detail elsewhere (Caplan, 1992. See especially pp. 403 -441).

Through a series of deletions, additions, and merging of tests which had previously been administered separately, the version of the battery which was used in this study consisted of 23 distinct subtests. In addition, Bub and I added 15 short subtests designed to test semantic access, as we had a particular interest in examining the differences in semantic access between PPA patients and DAT patients. However, those 15 tests are compiled into a single score for the purposes here, leaving a total of 24 subtests for consideration. This set of subtests was re-organized (during the development of the result diagrams which are described below) after all the data had been collected. This reorganization was necessitated by the fact that the original organization made it difficult to compare test results between modalities or in terms of Howard and Franklin's (1988) model of the functional organization of the language system, which will be discussed at the end of this chapter. The original organization also did not allow for easy understanding of the mediating role played by either semantic access or word type in language processes. In this thesis, the 24 subtests are presented in the conceptual framework which was developed after the data had been collected.

In this section both this interpretative framework and the test set itself are described, and the changes which were made to subtests for this study are documented. Where appropriate, results of an item analysis of the Boston norms are also presented. These results suggest several modifications that need to be made to the battery for future work. Detailed descriptions of how each test item was constructed are not included, since such descriptions has been published elsewhere (Caplan, 1992). Explanations of the method of administering the tests will be deferred until the final section of this chapter, which deals more precisely with the methodology of this study.

Problems With The Norming Procedure

Before the subtests themselves are described, four problems relating to the documentation of normal performance on the battery need to be pointed out.

The first problem is that practicality necessitated that the data collection proceed simultaneously with the lengthy process of analyzing the normal data which had been collected in Boston. (The norming procedure was identical to the testing procedure which is specified later in this chapter). As a result, information which was gathered from the analysis of the norms could not be used to improve the battery prior to its use in this study. This problem places me in the unpleasant position of having to undertake as part of the work presented here a formal critical analysis of the very tools I have relied upon for the study, and therefore to acknowledge some weaknesses in the work. Fortunately, most of these weaknesses can be overcome, and the few that cannot do not have devastating implications.

The second problem is that certain stimuli in a few of the tests were extended, or had items altered or replaced, after the norms had been collected but before the patients in this study were tested. Such changes were usually necessitated by the fact that the original items did not always control properly for the effect of lexical characteristics or by the recognition that the number of items in the test was too low. A few were changed for other pragmatic reasons. For example, in one case, the loss of resolution which had resulted from digitizing a visual stimulus made the item utterly incomprehensible. As a result, norms on some tests are not item-specific, but only class-specific: that is, we know how normals performed on the same class of stimuli presented in the same fashion, but we cannot be certain that the norms are applicable to the specific items in our version of this test. While this is clearly undesirable, it is not a fatal flaw in this study. With only a few exceptions (which are flagged in the test descriptions which follow), normals tended to perform near ceiling on the tests in the battery. It is thus extremely unlikely that normals who were able to perform almost perfectly in simple language tests on one set of stimuli would perform much worse on a very similar set of stimuli which differed mainly in being more carefully controlled for lexical characteristics.

The third problem which must be noted is that, because the norms were collected independently of this study and because the subjects of this study were not easy to find and thus could not be carefully selected, no attempt was made to match the normal control group with the clinical groups on any subject characteristics (and, indeed, the characteristics of the control group were not made available to me). While this is clearly not the way one would ideally like to collect data on normal performance, it must also be recognized that such this weakness, and even more glaring oversights of a similar nature, are extremely common in the field of clinical neuropsychology. It is extremely difficult to avoid such problems when one is working with human patients with rare conditions, or when one is relying upon tests which are expensive to administer and norm. It is difficult for practical reasons to eliminate this weakness within a single study, but it is possible for the field as a whole to compensate for it by insisting upon replication of findings across multiple patients (as discussed in Chapter 1) and across testing instruments.

The fourth and final problem with the norming procedure is that some of the normal result files were saved in a format which did not allow all of the data to be recovered. Consequently, it was not possible to include all results files from every normal subject in every calculation of normal performance. However, there is no reason to believe that this problem is in any way systematically related to normal performance, so it will not have biased the results. When result files did not include all the data necessary, they were eliminated from the calculation of normal results. This accounts for the variation in the number of subjects used to norm each test, since in actuality the same pool was used for every test.

Because the norms are currently unpublished and are not owned by me, they are not presented explicitly in this section unless they present problematic implications according to the following criteria: either the normal average score is below 90%, or the normal standard deviation is greater than 10%, or the normal average minus two normal standard deviations (that is, the lower bound of the 95% confidence interval) is less than what would be expected by chance, when such a chance score is calculable. The first two criteria rely upon ad hoc values which reflect a reasonable assumption about normal performance on language tests: that normals should be able to perform quite well on them. In this case, 'performing quite well' is defined as having a lower bound on the 95% confidence interval of the test which is equal to a score of 70%.

Many of the tests are controlled on several factors. In order to allow for the possibility of statistically-significant discriminative power, the analyses in this thesis ignore any factors (with a few necessary exceptions which will be noted) which define a subset of less than eight items.

In the item analyses which follow, items were flagged as 'bad' only if the average score for normal subjects on that particular item was at least two standard deviations below the normal average score across all items, *and* if the percentage of normal subjects who failed the item was greater than 10%. The latter criterion serves to prevent items from being flagged as poor simply because the normal average score was very near ceiling.

A.) Tests Of Word Access
i.) Non-semantically mediated
1.) Phoneme Discrimination (PD)
2.) Auditory Lexical Decision (ALD)
3.) Written lexical Decision (WLD)
4.) Reading (READ)
5.) Repetition (REP)
ii.) Semantically-mediated
Word Comprehension
6.) Auditory Word Picture Matching (AWP)
7.) Written Word Picture Matching (WWP)
8.) Semantic Access (SEMANTIC)
Word Production
9.) Oral Naming (O-NAME)
10.) Written Naming (W-NAME)
B.) Tests Of Word Category Access
i.) Abstract Words
11.) Auditory Comprehension Of Abstract Words (A-ABS)
12.) Written Comprehension Of Abstract Words (W-ABS)
ii.) Affixed Words
13.) Oral Production Of Affixed Words (O-PROD-AFF)
14.) Written Production Of Affixed Words (W-PROD-AFF)
15.) Auditory Lexical Decision Of Affixed Words (ALD-AFF)
16.) Written Lexical Decision Of Affixed Words (WLD-AFF)
17.) Auditory Synonym Judgment Of Affixed Words (A-SYN)
18.) Written Synonym Judgment Of Affixed Words (W-SYN)
19.) Auditory Word Picture Matching Of Affixed Words (AWP-AFF)
20.) Written Word Picture Matching Of Affixed Words (WWP-AFF)
C.) Tests Of Sentence-Level Processing
i.) Sentence Comprehension
21.) Auditory Comprehension Of Sentences (A-COMP-SEN)
22.) Written Comprehension Of Sentences (W-COMP-SEN)
ii.) Sentence Production
23.) Oral Production Of Sentences (O-PROD-SEN)
24.) Written Production Of Sentences (W-PROD-SEN)

Table 2.1
The 24 subtests of the PAL battery used in this thesis

The 24 subtests have been divided here into three main conceptual sets of tests (see Table 2.1): a set testing general word access, a set examining performance at processing the meaning of words from specific word categories, and a set designed to measure sentence-level processing. They shall be examined here in this order. Details of presentation (e.g. font size, presentation rate, digitization rate) for all 24 tests are presented in the methodology section in the second half of this chapter.

A.) Tests Of Word Access

The tests of word access are themselves divided into two groups, which have been titled 'Nonsemantically-mediated' and 'Semantically-mediated'. The title is somewhat inaccurate, since both the nonsemantically-mediated tests and the semantically-mediated tests require the subject to have access to some semantic information about the words themselves. In this strict sense the only genuinely nonsemantically-mediated test is phoneme discrimination, since that is the only test that does not require that one even be able to recognize words. (Unlike phoneme discrimination, repetition is, perhaps surprisingly, quite sensitive to a patient's ability to recognize words.) The names of the two groups of subtests reflect the fact the tests which compose them differ in the extent to which they require one to have access to semantic information about the word's *referent*. The distinction will be clear to anyone who has ever played Scrabble with an expert Scrabble player: it is common for such players to put down memorized words which they are sure are real words but whose meaning they do not know. (My own experience of being married to such an expert suggests that in such cases the word will invariably turn out to refer to an obscure African ruler!) Most of us are able to read, repeat, and even recognize and spell words whose meaning we do not know. However, we cannot, of course, reasonably expect to name items whose name is a word whose meaning we do not know. In cases of brain damage, more subtle effects of semantic access are possible, as shall be seen in some of the case studies presented in Chapters 4 and 5.

i.) Tests of nonsemantically-mediated word access

There are five tests in the nonsemantically-mediated word access category: phoneme discrimination, auditory and written lexical decision, reading of words and nonwords, and repetition of words and nonwords.

In order to make the symmetry between modalities complete, two other tests should have been added to this section: copying written words (the written analog of repetition), and writing to dictation (the 'reverse' of reading aloud, inasmuch as it maps from oral input to orthographic output). The latter test was part of the original PAL battery, but was dropped from the version used in this study because a new version could not be prepared in time.

1.) Phoneme Discrimination

The ability to discriminate between phonemes is assessed by presenting the subject with 40 recorded pairs of monosyllabic nonwords (e.g. doss/doz). The subject is asked to decide whether he has heard the same stimulus two times, or two different stimuli.

This test was normed on 54 subjects. The norms were within the range defined earlier as acceptable.

The following four items had average scores which were more than 2 standard deviations below the normal average score: doss/doz, kest/keft, flom/slom, and pebs/peds. Notice that three of the bad items have differing fricative sounds (that is, 's', 'z', or 'f' sounds), which are notoriously difficult to record, and easily degraded in low fidelity recordings. The remedy for the bad items is thus probably to re-record all the stimuli with a higher fidelity.

2.) Auditory Lexical Decision

Subjects listened to either a word or nonword. They were asked to say whether each stimulus they heard was a real word or a made up word.

The subtest was normed on 56 subjects. The normal averages and standard deviations were well within the acceptable range. The original PAL subtest contained 40 concrete nouns, balanced on length and frequency, and 40 nonwords foils. Bub and I expanded the subtest by adding 22 similarly balanced abstract words.

The item analysis found two items with average scores more than 2 normal standard deviations below the overall normal average score: the nonword 'shess' and the word 'moth'. The first item is probably too close to the word 'chess', and needs to be replaced for that reason. The problem with the word 'moth' is not clear. However, with a normal error rate of 43% on that item, normal performance at recognizing this word was not significantly different from chance ($\chi^2(1) = 1.1$; $P > 0.05$).

3.) Written Lexical Decision

Subjects were presented visually with either a word or nonword. They were asked to say whether the stimuli they saw was a real word or a made up word.

The test was normed on 54 subjects. Normal averages and standard deviations were again well within the acceptable range. The original test contained 10 concrete nouns and 10 abstract nouns, balanced for orthographic regularity, and 20 nonword foils. We added six items to each of the two noun categories, and changed other stimuli in order to balance for length and frequency as well as regularity. We also added 12 new nonwords, and balanced all 32 nonwords on length.

The item analysis of the stimuli in this subtest did not reveal any bad items.

4.) Reading

Subjects were presented visually with either a word or an orthographically-regular nonword, and were asked to read the item aloud. We altered this test from its original form to allow the tester to keep a record of the subject's actual response as well as a score for each item, in order to facilitate the classification of errors.

We used the original version of this test, which consists of 27 words balanced on frequency, length, orthographic regularity, and imageability (abstractness), and added the nonword reading task which had originally been administered separately. The nonword reading task consists of 32 nonwords balanced on two levels (simple or complex) of orthographic and phonological complexity.

Normal data on this test was collected for only 14 subjects. However, norms on those 14 subjects were, as one might reasonably expect with such an easy task, extremely high.

There were only two items on which more than one subject made an error. In both those cases only two subjects made errors. However, because of the low number of control subjects, these items do in fact meet the criteria for identifying bad items. They are the nonwords 'ras' and 'snike'. The first of these items is clearly ambiguous. The rules of orthography suggest that it should be pronounced 'rass', but since the only short words in English which end with 'as' are the irregular words 'as' and 'has', one is tempted to pronounce it as 'razz'. Such ambiguity is not intended, so that item should certainly be replaced. It is more difficult to speculate why 'snike' might have tripped up two subjects. There is no reason to expect the stimuli will be an outlier with a larger group of normal subjects.

5.) Word And Nonword Repetition

Subjects heard either a word or nonword, and were asked to repeat the item exactly as they heard it. Like the reading test, this test was altered so as to allow the tester to enter the subject's response for latter classification of errors.

The original subtest contained 20 words balanced on frequency and orthographic structure (10 CVC versus 5 CCVCC versus 5 complex), and 20 nonwords balanced in the same way on orthographic structure only. We added 10 more words, and changed some of the extant stimuli so that the words were also balanced evenly on imageability.

Although normal performance (with 52 subjects) in repeating words is very high, the normal average in repeating nonwords was only 84%. With a standard deviation of 12%, the normal individual's 2 SD cut-off point for nonword repetition is only 60%. Since repeating nonwords is normally considered an extremely simple task, we must assume that recording fidelity caused the problem.

Three items, all nonwords, were below the population's 2 SD cut-off point, having been correctly repeated by less than 60% of the normals. The three problematic stimuli are 'heen', 'temasone', and 'nid'. The third one was often lexicalized to 'mid' by the subjects.

ii.) Tests of semantically-mediated word access

There are five tests in the semantically-mediated word access category, which may themselves be conceived of as forming two subcategories, one testing comprehension and the other testing production of such words.

Tests Of Semantically-Mediated Word Comprehension

The three subtests which test production of semantically-mediated words are auditory and written word-picture matching, and the test of semantic access.

When we began collecting data, we had included a fourth test in this category, picture homophone matching. This is an important test, because it allows one to test a case which is not tested otherwise: namely, that the patient has internal access to lexical phonological representations, but is not able to produce phonological output. The test included 20 stimuli containing two pictures. The patient was asked to signal if the two pictures had homophonic names, as, for example, of we showed a picture of a baseball cap and a bottle cap. The negative trials included pictures with names which differed by only one feature: for example, a picture of a stove and a stone.

We stopped administering this test because of our judgment that it was not sufficiently constrained to test what it was supposed to be testing. We noticed that patients did not always name the pictures using the homophonic word. A bottle cap and a baseball cap are not homophones if one calls the baseball cap 'a hat', or even if one considers the modifier nouns 'bottle' and 'baseball' to be part of the name. In this case, our clinical judgment was shown to be justified by formal reasoning. Normal average performance at judging the pictures which were intended to have the same name was 89%, with a standard deviation of 13%. The 2 SD cut-off point is thus 63%. Since the test included 16 such pictures, the cut-off point is equivalent to 10 correct items, which is unfortunately not different from what one can expect by chance alone on a binary-choice test ($\chi^2(1) = 1.0$; $P > 0.05$). The test thus does not have sufficient power to discriminate normal performance at identifying homophones from chance performance.

6.) Auditory word-picture matching

In the auditory word-picture matching test, subjects see a black and white line drawing of a common animal, fruit, vegetable, or inorganic object upon the screen, and simultaneously hear a word. They are asked to tell the examiner if the word they heard corresponds to the name of the picture.

This was one of the few tests whose format we changed. In the original test, subjects saw two pictures on the screen and had to decide which one was the correct picture to go with the word. We felt this was a poor test, since subjects only had to decide which one of the two drawings was the best match for the word, rather than having to decide if there was a match at all- and thus could often answer correctly even with limited knowledge. An example will make the problem clear: Imagine being asked about the flag of the new country of Georgia. Most of us have been left behind by the many changes in the political situation in the old Soviet Union, and could neither describe the flag, nor decide whether an arbitrary flag was the right one or not. However, we could certainly choose the right flag from a two-choice pairing if the foil was a flag we recognized. Using the old scheme, a person would get 'full marks' from this latter choice, even though they actually knew nothing whatsoever about the target- the flag of Georgia. Under the new scheme, a person would get only half marks (for knowing that the foil was not the flag of Georgia) if they had some knowledge about the foil but none about the target. Similarly, a person who knows nothing about either flag had a 50% chance of getting 'full marks' despite his lack of knowledge under the old scheme, but only a 25% chance of getting full marks under the new scheme, since he now needs to guess correctly between two choices two times. We kept the same 32 stimuli as in the original test, but presented them as 64 one-choice match decisions.

Unfortunately, most common animals and objects have names which appear with low frequency on the word corpus which is used to assess frequency. As a result, it is very difficult to come up with a test of word picture matching which is perfectly matched on frequency. Both the auditory and written word picture matching tests in the PAL contain only 8 pictures with high frequency names, compared to 56 with low frequency names. This means that significant frequency dissociations on these tests, especially in favour of objects with low frequency names, must be interpreted cautiously.

The normal average score obtained by 50 subjects on the original version of this test (which we may suppose to slightly over-estimate performance on our slightly more difficult version) was high, with a small standard deviation. Nevertheless, there were two bad items, each correctly matched by only 60% of the subjects, a rate that is not significantly different from what we could expect by chance ($\chi^2(1) = 2$; $P > 0.05$). The two items were a syringe, paired with a thermometer as a foil, and a picture of celery, paired with lettuce as a foil. The two sets of pictures are shown together in Figure 2.1.

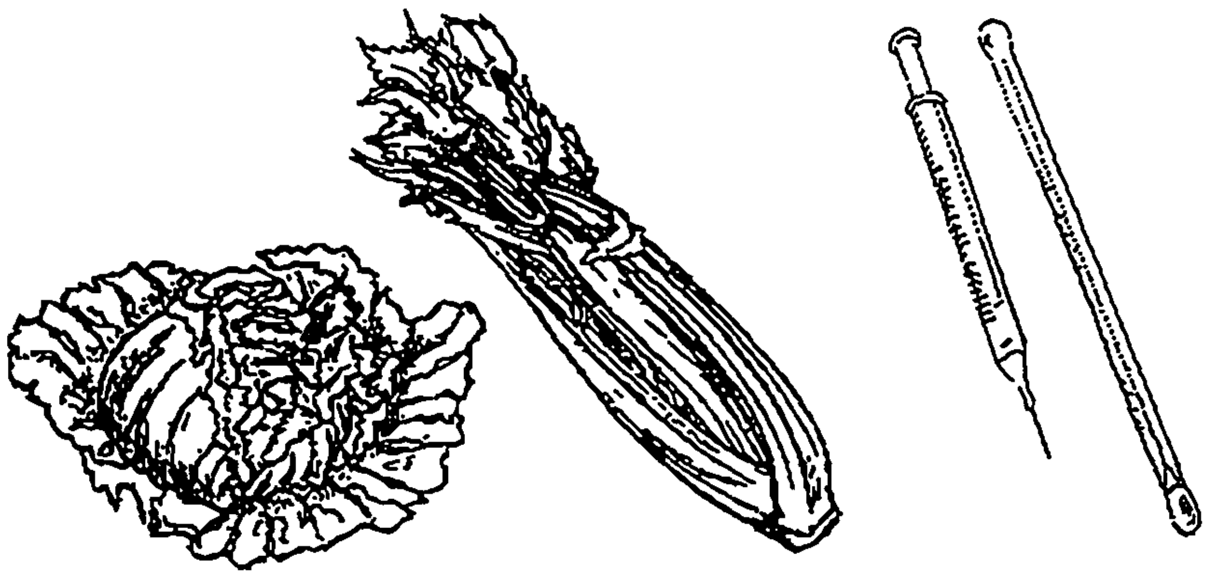


Figure 2.1:
Bad Items From The Auditory Word-Picture Matching Test
Celery (target), with lettuce foil
Syringe (target), with thermometer foil

7.) Written word-picture matching

The structure of the written word-picture matching test is identical to the oral test in every way, except that the subject sees a name printed on the screen rather than hearing it pronounced aloud. The stimuli were drawn from the same three classes (fruits and vegetables, animals, and inorganic objects), though the stimuli were not the same as the stimuli in the auditory test.

We changed this test in exactly the same manner that we changed the auditory version: i.e. by turning a forced choice task into a decision task.

Normal performance (from 52 subjects) on this test was within the acceptable range defined earlier.

A single item fell into the range defined for identifying bad items: a picture of a moose which was paired with a picture of an elk as a foil. If only one subject more had answered the item correctly, however, the average score for that item would not have met the first criteria for identifying bad items: namely, that at least 10% of normal subjects fail on that item.

8.) Semantic Access

Subjects were presented with 15 subtests designed to assess their access to knowledge about animals, a category chosen because knowledge about animals (or, more generally, living things) is known to be particularly fragile (Farah et al, 1991). All tests had the same format: either the written name or a picture of a target animal was presented in the middle of the screen, and subjects were asked to match that name or picture to one of four stimuli around the animal, by pointing to the correct match. The four choices and their relation to the stimuli varied by subtest. Each subtest consisted of between 11 and 25 trials. A summary of the 15 subtests is presented in Table 2.2.

1.) ENV-P: Match animal pictures to their environment
2.) ENV-W: Match animal names to their environment
3.) FEET-P: Match animal pictures to their feet
4.) FEET-W: Match animal names to their feet
5.) FOOD-P: Match animal pictures to their food
6.) FOOD-W: Match animal names to their food
7.) HEIGHT-P: Match animal pictures to objects of same height
8.) HEIGHT-W: Match animal names to objects of same height
9.) HORNS-W: Match animal names to appropriate picture varying on horns and teeth
10.) LENGTH-P: Match animal pictures to objects of same length
11.) LENGTH-W: Match animal names to objects of same length
12.) MARK-P: Match animal pictures to their skin markings
13.) MARK-W: Match animal names to their skin markings
14.) TAILS-P: Match animal pictures to their tails
15.) TAILS-W: Match animal names to their tails

Table 2.2:
Name and brief description of the 15 tasks which make up the semantics subtest

In order to keep the information about semantic access manageable in the context of this study of general language deficits, results from all fifteen tests were averaged together get a global score for semantic access. Average scores for access from words versus pictures and for domestic versus foreign animals were also calculated.

Normal performance on the semantics tests (with 14 elderly subjects) were gathered in Montreal some years ago. Normals found the tests quite difficult. The average score over all tests was below our criterion, at 89% with a standard deviation of 12%, giving a 2 SD cut-off point of 65%. However, note that since this set of tests has a chance score of just 25%, this low score is well above chance ($\chi^2(1) = 128$; $P < 0.01$).

Because we averaged together results from all 15 subtests designed to assess access to semantic knowledge, the item analysis of the individual tests is not directly relevant and so is not presented here.

Tests Of Semantically-Mediated Word Production

The two subtests which test production of semantically-mediated words are the picture-naming tests in the written and spoken modalities.

9.) Oral Naming

In the oral naming task, subjects are presented with a line drawing of a common animal, fruit, vegetable, or object, and must tell the examiner the name of the picture. As well as being balanced on semantic category, the 32 stimuli are balanced on name length and frequency.

Overall norms on this test (collected from 67 normal subjects) were within the acceptable range.

There were three bad items, all drawn from the category of animals: 'goat', 'buffalo', and 'leopard'. It would be easier to understand these errors if we had a record of what the normal subjects answered. Since none of these animals is obscure, one suspects that the subjects were marked wrong (as they should have been) for producing close semantic neighbours such as 'sheep', 'ox', and 'panther'.

10.) Written Naming

As in the oral naming task, subjects are presented on this test with a line drawing of a common animal, fruit, vegetable, or object, whose name they must produce, in this case by writing it. As well as being balanced on semantic category, the 32 stimuli in this test are balanced on name length and frequency.

Normal performance on this test (assessed on 61 subjects) is well within the range defined as acceptable.

Average scores for two items fell outside the cut-off point: 'broccoli' and 'chisel'. Since these are very common items which we would intuitively expect normal subjects to be able to easily recognize, the poor performance must presumably be due to the quality of the pictures. These have been reproduced in Figure 2.2. Again, a record of the normal responses would probably make the problem with the stimuli clear.

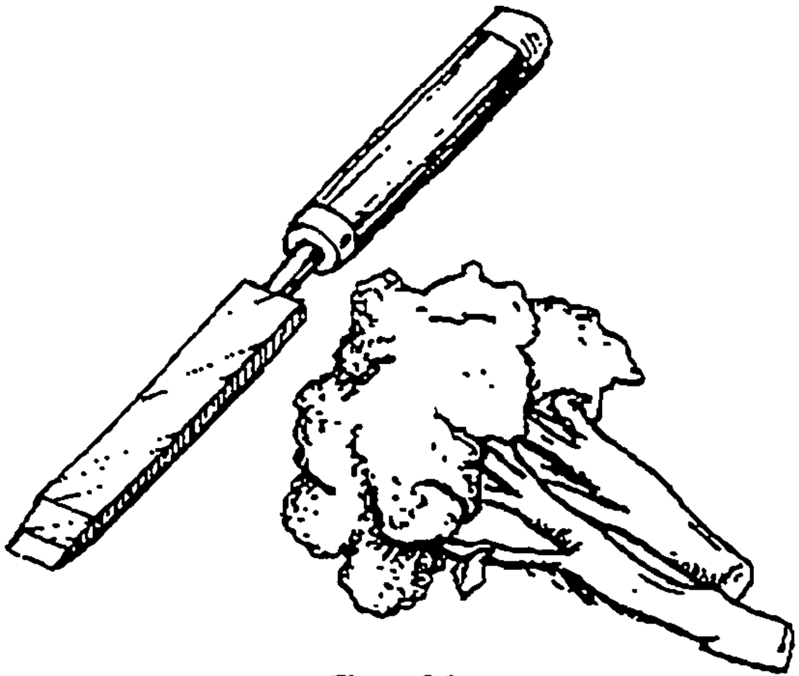


Figure 2.2:
Bad Items From The Written Naming Test
Broccoli & Chisel

B) Tests Of Word Category Access

The PAL battery looks at two word types in particular detail: abstract words, and affixed words.

Tests Of Abstract Word Comprehension

Two tests examine abstract word comprehension in the written and spoken modalities. In both cases, subjects are presented with an abstract target word, and two possible synonyms. They must choose which of two possible synonyms is closest in meaning to the target, by stating their choice. Because of the great difficulty in matching abstract words, the two tests use the same 20 stimuli to make cross-modal comparisons possible. The abstract target words are all short (mono- and bi-syllabic) but otherwise are not controlled for stimulus characteristics such as frequency.

There are no tests of abstract word production, for the simple reason that it is not at all clear how one might construct such a test.

11.) Auditory Comprehension Of Abstract Words

The stimuli in this test are presented in the auditory modality, with a beep separating the target from the synonym choices.

Average normal performance (by 54 subjects) on this test was within the acceptable range.

No items had an average score which fell outside of the 2 SD cut-off point.

12.) Written Comprehension Of Abstract Words

The stimuli in this test are presented visually, with the target on a separate line from the synonym choices.

Average normal performance (by 55 subjects) on this test was well within the acceptable range.

A single item fell outside of the range we have defined as acceptable, having been correctly answered by 81% of the subjects (significantly above chance: $\chi^2(1) = 21$; $P < 0.01$). The target was 'sincere', paired with 'honest' or 'ready'. It is puzzling why this item should have been answered so much more poorly in the written than the auditory version of the abstract word comprehension test.

Tests Of Affixed Word Production And Comprehension

The tests of morphological processing (access to derivations and inflections) are categorized into two sections, one testing production and the other testing comprehension.

Tests Of Production Of Affixed Words

The two tests of affixed word production use the same stimuli. In each case, the subject is presented with the base form of a word, followed by a sentence with one word missing. The subject must produce the affixed form of the initial word which correctly completes the sentence. An example of one stimuli is 'SING: Last night the choir ____ beautifully'.

Both derivational affixation (affixes which create a new category of word from a common word stem: i.e. the affix in 'ness' in the word 'happiness') and inflectional affixation (affixes which do not create a new category of word: i.e. the affix 's' in the word 'sings') are used. Within derivational affixation, the ability to produce both word-boundary affixes (or level 2 affixes, which leave the stem unchanged) and formative-boundary affixes (or level 1 affixes, which change the stem) is tested.

13.) Oral Production Of Affixed Words

In the oral test, the subject hears the base form of the word, followed after a brief pause by the sentence, with a beep to indicate where the word is missing. The subject is asked to complete the sentence by stating the correct form of the target word.

Norms are available for only 13 subjects. The average score and standard deviation were within the acceptable range.

A single item fell outside of the acceptable range. The stem word was 'receive', with the cue sentence 'The boss was very ________ to the request for a raise.' Unfortunately, the low number of normal subjects makes it impossible to know if this item is truly problematic: only 2 subjects made an error in responding on this item. Note, however, that the same item posed no problem in the written version - and we would not expect normals (by virtue of the very fact of their normality) to show a dissociation between written and spoken morphological marking, particularly a dissociation which favours written output.

14.) Written Production Of Affixed Words

In the written test, the stimuli are presented on the screen, with the base form of the word presented above the sentence. The subject must complete the sentence by writing down the correctly affixed form of the target word.

Average normal scores and normal standard deviations on this test (with 63 subjects) are within the acceptable range.

There are no bad items in this test.

Tests Of Comprehension Of Affixed Words

The PAL battery includes six subtests which assess comprehension of word morphology. These subtests can be organized in the same manner as the more general tests of lexical access, into a set of subtests which are nonsemantically mediated (in the sense explained at the beginning of this section) and a set which is semantically mediated.

Tests of Nonsemantically Mediated Affixed Word Comprehension

There are two non-semantically mediated tests of affixed word comprehension: written and oral lexical decision of affixed words. In both tests, the subject is presented with either an affixed word, or a plausible nonword, and must say whether or not the stimuli is a real word. As in the subtests of affixed word production which were described above, the words in the comprehension tests vary on type of affixation (derivational and inflectional), and type of inflection (word-boundary and formative-boundary). The nonwords are constructed by affixing a real root with a real affix which is not usually affixed to that root. For example, the test includes the nonword 'happiment', formed by affixing 'ment' to the stem word 'happy'.

15.) Oral Lexical Decision Of Affixed Words

This test is identical to the written lexical decision subtest described above, the only difference being the nature of the stimuli and foils. This structural identity suggests a small improvement in the test protocol with respect to these two tests: there is no reason why the affixed words and foils should not be presented to the subject at the same time as the rest of the lexical decision words, since the test structures are identical. The separation of word types that are of interest to the researcher or clinician can be done at analysis time. This would speed up the testing slightly, because it would save the small amount of time that is spent getting organized for the next subtest, and because patients often measure the amount of time they will devote per session in terms of how many tests they have completed.

The norms of this test (calculated from 28 normal subjects) were within the acceptable range.

The rules of English word affixation are extremely productive- that is, they allow one to coin acceptable terms rather easily (Pinker, 1994). In light of this productivity, it should not be surprising to find that there might be some ambiguous items in this test. The surprising finding, however, is that two of the four items that failed to meet the criteria for acceptability were real words: 'supposes' and the adjectival reading of the word 'animate'. It is difficult to speculate why these words should cause normals any difficulty. The other two bad items were the nonwords 'caming' and 'momental'. The worst of these words ('caming', mis-identified by 25% of the normals) was still correctly identified at a rate which was significantly better than chance ($\chi^2(1) = 7$; $P < 0.01$).

16.) Written Lexical Decision Of Affixed Words

This test was identical in form to the subtest of written lexical decision described above, and used the same stimuli as the subtest of auditory written lexical decision of affixed words.

Normal performance on this test (as measured with 61 normal subjects) is within the acceptable range.

Only two terms in this test caused trouble for a substantial number of the normal subjects: the nonwords 'amateurist' and 'disappearable'. Unlike many of the nonwords in the test (such as 'distanter' and 'caming') these two words do not violate any rules of affixation. You will probably have no trouble understanding what the following two sentences mean: 'A well-known stage magician has proven that even a Boeing 747 jet is disappearable' and (rather more bizarrely) 'My cousin is always placing ads looking for inept beginners to use as subjects for his studies, because he is a professional amateurist'. If you don't have any trouble understanding those two sentences, then you will have to agree that the two neologisms they contain should be removed from this subtest (and perhaps considered for inclusion in the next edition of the Oxford English Dictionary?)

Tests of Semantically Mediated Affixed Word Comprehension

There are two tests of affixed word comprehension in each of the written and auditory modalities which require the subject to have access to the meaning of the words: synonym judgment of derived words and word-picture matching of derived words.

17.) Auditory Synonym Judgment Of Affixed Words

This test is identical in form to the auditory abstract word judgment test described above, differing only in that it uses affixed words as targets. As in that test, subjects are presented with three stimuli. They must indicate verbally which one of the two final stimuli is most similar in meaning to the first stimulus. The 20 stimuli in this test could be presented together with those stimuli to streamline the administration of the battery.

Sixty normal subjects achieved an average score of 88% on this test. With a standard deviation of 11%, the cut-off point on this test is only 66%, which is not significantly better than the score one can expect by chance on a 20 item binary choice test ($\chi^2(1) = 1.8$; $P > 0.05$). This greatly limits the use of the test, since it cannot be used to identify poor performance, but only spared performance. Even that can only be identified with very weak certainty, since a ceiling effect prevents a person from scoring more than one standard deviation above the normal average.

Because normal subjects had difficulty with many items in this test, only one item fell outside the cut-off point for identifying poor items. The stimulus item was the word 'judgment', with the choices 'ruling' and 'ruled'. Removing this item does not change the undesirable characteristics of the normal response distribution. Accordingly, results from this test were necessarily removed from consideration in this study.

18.) Written Synonym Judgment Of Affixed Words

This test is identical in structure to the written abstract word judgment test described above, but uses only affixed words as targets. It uses the same 20 stimuli as the auditory version of the affixed word synonym judgment test.

Performance was much better than performance on the auditory version of the same test. Scores from 62 normal subjects on this test had an average and standard deviation which were above the cut-off points for identifying problematic patterns of normal performance.

The only item which fell below the cut-off point for identifying bad items was the same item which had been identified in the auditory version of this test. This cross-modal agreement adds further weight to the suggestion that the item does indeed need to be changed.

19.) Auditory Word Picture Matching Of Affixed Words

In this test, subjects saw two pictures on the screen, while they heard an affixed word. Their task was to choose which picture was best described by the affixed word. The test includes 20 items. We were not able to improve the administration of this test as we had improved the administration of the general word picture matching tests (by matching a single word with a single test), as the pictures on this test differ in subtle ways, and thus can only be disambiguated when they are seen together.

The test was normalized on 39 normals. Their average score was 88%, with a standard deviation of 8%. The cut-off point for identifying normal performance is thus 72%, which is again not high enough to allow normal performance to be distinguished from chance performance on a binary-choice test with only 20 items ($\chi^2(1) = 3.2$; $P > 0.05$).

When this test is analyzed by item, the standard deviation is so high (22%) that no item meets the resultantly stringent criteria for identifying bad items. However, in this case we do not need statistics to help us: the worst of the bad items are clearly apparent upon visual inspection of the results. Seventeen items were answered correctly by at least 77% of the normals. The remaining three items are the only three items which were answered correctly at a rate that was not significantly better than what chance would predict, being answered correctly by 52%, 46% and just 27% of the normals. This latter item was also the only item in any one of the 24 tests which was answered incorrectly by a number of subjects differing significantly from chance, but in the wrong direction! The stimulus was the word 'floral', paired with a picture of a woman arranging flowers in a flower shop, and a picture of wall-paper which is supposed to have flowers on it. The flowers on the wall are very difficult to see in the digitized image, so most subjects chose the picture of the woman arranging flowers. Since flower arrangements are sometimes called floral arrangements, the item is ambiguous, enough to lead normal subjects to make the wrong choice far more often than the right choice. The second worst item is the word 'raising', paired with a picture a man who has just raised some weights over his head, and a man who is in the process of raising the weights over his head. The distinction is obviously not clear to our normal subjects. The final bad item is the word 'reflector' paired with pictures of a bicycle reflector and of a woman looking into a hand mirror. The correct answer is supposed to be the bicycle reflector, but apparently normal subjects are equally happy to call a mirror a reflector.

If these three stimuli were replaced with good items, or even simply deleted, this test would be quite acceptable. Across the 17 items which remain when they are deleted, the average score was 95%, with a standard deviation of 6%, which is well within the level we have defined as acceptable. The 2 SD cut-off point specified by these figures differs significantly from a chance level ($\chi^2(1) = 7.1.$; $P < 0.01$). In view of these findings, all patients in this study have been scored on the abbreviated 17-item subtest.

20.) Written Word Picture Matching Of Affixed Words

This subtest is identical to the test just described, except that the target is presented visually on the screen. Subjects saw two pictures and an affixed word on the screen, and had to choose which picture was best described by the affixed word. The visual stimuli are the same as the ones used in the auditory test, but the lexical stimuli are different- thus the bad items on the auditory version will not necessarily be problematic in the written version.

The test was normed on 52 normal subjects. The average scores and standard deviations were within the acceptable range.

Two items fell into the category of bad items. One was the word 'jumped', paired with pictures of a boy who has just jumped or is about to jump from a diving board. The other is the word 'sculpture', paired with pictures of a sculpture and of a sculptress in the process of carving a stone.

This latter item is overtly ambiguous, since both pictures contain sculptures.

C.) Tests Of Sentence Comprehension & Production

The battery includes tests of sentence comprehension and production in both the written and spoken modalities.

Tests Of Sentence Comprehension

The tests of sentence comprehension are composed of two subtests which are conceptually distinct. Both subtests use a sentence picture matching technique: the patient hears a spoken sentence or sees a written sentence, and must choose which of two pictures correctly matches the sentence.

Twenty of the 40 stimuli in the subtest are intended to assess performance at comprehending semantically-constrained sentences: that is, sentences in which the subject and object cannot be confused due to real-world limitations. For example, a man may wax a car, but a car cannot wax a man, so the sentence 'The man is waxing the car' is semantically constrained.

These items are intended to test understanding of word roles. They vary with respect to the type of word they test (by varying the foil): ten items test verbs (for example 'The lawn was mowed by the boy' versus 'The lawn was raked by the boy'), five items test prepositions (for example 'The boy is walking to the house' versus 'The boy is walking from the house') and five items test the role of particles (for example, 'The man is hanging up the shirts' versus 'The man is hanging out the shirts'). The sentences testing verbs are divided into five sentences which test active verbs (for example, 'The boy is throwing the ball') and five which test for passive verbs (for example 'The wagon was pushed by the girl.'). Because there are so few exemplars of each word role, in this thesis patient performance is analyzed only across all 20 constrained sentences.

The remaining 20 stimuli are semantically-reversible: that is, the role of the subject is not constrained by real-world limitations. For example, a man may hug a woman just as easily as a woman may hug a man, so the sentence 'The woman hugs the man' is not semantically constrained. These items are intended to test syntactic comprehension. The 20 sentences test four different sentence structures: active ('The man hugged the woman'), passive ('The man was tickled by the girl'), dative-passive ('The pig was pushed to the goat by the cow'), and subject-object relative ('The sailor that the soldier pushed hit the policeman'). In this thesis performance is analyzed only across all 20 semantically-reversible sentences.

21) Auditory Comprehension Of Sentences

Normal performance (by 61 subjects) on the constrained sentences as a whole was within the acceptable range. However, the normal average score on the 20 reversible sentences was just below our identified cut-off point, at 89% (SD = 7%). Those items still have a cut-off point which is significantly different from chance ($\chi^2(1) = 5.$; $P < 0.05$).

Despite a large variance in normal performance across the items, two items met the criteria for identification as bad items, one having been answered correctly by only 58% and the other by only 15% of normals.

The first problematic sentence was 'The robber held up the man', paired with pictures of a man in a mask holding back a man or holding him up at gun point. This sentence is problematic in a number of ways. The phrase 'held up' is ambiguous, since it can indeed be used to mean 'held back', as in the sentence 'Sorry I'm late: I was held up by traffic'. Moreover, the sentence is not clearly semantically constrained, though it is in that category, since a man can indeed hold back a robber as easily as a robber can hold back a man.

The second problematic sentence was the semantically-reversible subject-object relative sentence 'The sailor that the soldier pushed hit the policeman', paired with pictures of a soldier hitting a sailor, followed by a picture of either the soldier or the sailor hitting a policeman. Perhaps the images are not detailed enough to make clear that the characters in the two scenes are intended to be the same character.

22) Written Comprehension Of Sentences

Normal performance (by 62 subjects) on the written constrained sentence comprehension was within the acceptable range. However, as in the auditory modality, the normal average score on the 20 reversible sentences was below the identified cut-off point, at 87% (SD = 9%). This distribution is not sufficient to allow subnormal performance to be reliably distinguished from chance performance ($\chi^2(1) = 3.2$; $P > 0.05$). For technical reasons, it is difficult to re-analyze normal performance on these items with the bad items removed, in order to see if such changes could salvage the subtest's ability to discriminate normal from chance performance. I have therefore removed the reversible sentences from further consideration.

Although the normal standard deviation on the item analysis is again quite large, three items were nevertheless identified as bad items.

Only one of the three bad items (answered correctly by 66% of the normals) was from the semantically-constrained set. The sentence was 'The cat is in the tree'. However, due to loss of resolution during digitization, no cat is visible in the tree in either picture, making the item very difficult indeed. Twenty-two percent of normal subjects either refused to answer the question, or indicated that there was no correct answer. We replaced this item on the version of the test that we used.

The other two bad items were both semantically-reversible subject-object relative sentences: 'The man that the woman followed kicked the boy' (answered correctly by only 71% of normals), and 'The dog that the cat bit chased the boy' (answered correctly by 68% of normals).

Problems With The Sentence Comprehension Tests

Although to do so I must ignore the decision I made to avoid analyzing item sets with less than eight items, it is important to point out that the weight of evidence makes clear that these tests of sentence comprehension are deeply flawed. These tests are intended to be analyzed by the four semantically-reversible subtypes and four semantically irreversible subtypes, as described above. However, it is not in fact possible to analyze the tests in this manner. To see why, let us consider first the format for testing comprehension of semantically-reversible subject-object relative sentences. Three of the ten such items appearing on the two comprehension tests were identified as bad items. Moreover, normal performance on all of the subject-object relative sentences is absurdly poor. The average normal score for such sentences on the written test was 77%, with a standard deviation of 21%, given a 95% confidence interval lower bound of only 35% on this binary-choice test. Average performance on the auditory version of the test was even worse: with an average of only 67% and a standard deviation of 19%, the 2 SD cut-off point is only 29%. It need hardly be pointed out that these distributions are not significantly different from the distribution predicted by chance, since, with only five items per set, a normal average of only a single error (that is, an average score of 80%) yields a score that is not significantly different from chance ($\chi^2(1.8) = 3.2$; $P > 0.05$). In order to be able to distinguish normal performance from chance performance with only five binary-choice items, the normal average score must round to 100%, which in practice means that the 2 SD cut-off point must be above 90%. This level of normal performance was not attained on any one of the eight sentence categories in each of the two sentence comprehension tests, so no meaningful analysis of performance by sentence type would be possible on either of these tests. The tests must either be extended, so that each sentence category is tested with many more items, or the items must be re-designed so that normals almost never make errors.

Sentence Production

The 23rd and 24th tests are the tests of oral and written sentence production. In these tests, subjects were shown a picture on screen and asked to produce a sentence in the appropriate modality which described it. In order to allow for comparisons between modalities, both tests used the same pictures. The productions were constrained in three ways. Subjects were given the verb which they were to use, which was printed in its infinitive form on the screen. Each picture had a dot beside two or more items, indicating that those items had to be mentioned explicitly. One of those items also had an arrow beside it, to indicate that the item had to be mentioned before any other item. This last constraint is necessary in order to force subjects to produce sentences with a dative-passive construction. Since the test was intended only to assess sentence construction, subjects were given as much help as they requested in naming objects.

The test included 25 sentences in five blocks by sentence type: active ('The boy is hugging the girl'), passive ('The girl was pushed by the boy'), relative clause ('The man carrying a newspaper is lifting a bag'), dative ('The girl is putting a book on the table'), and dative-passive ('The ball was thrown to the boy by the man'). Although normal data was collected for this test, differences in the way the tests were administered to the normals and to our own patient population made them inappropriate for our use. The test is thus scored only to identify dissociations between sentence types within a single patient, and to assess in a qualitative manner the patient's ability to produce sentences.

ii.) The PAL Battery Analysis System

The practical difficulty of scoring, cross-referencing, and interpreting the results from the 24 tests in the aphasia battery presented a considerable practical barrier to its use. In order to surmount this barrier, I developed a hybrid Common Lisp application which contains a production system: a computational system which relies upon self-contained data-sensitive rules (also called 'productions') rather than defined procedures to compute (see Buchanan & Shortliffe, 1984; Brownston et al, 1985). In this subsection, this analysis system is briefly described.

The system has the following functions:

- to analyze each of the raw result files along the relevant dimensions, as well as along all pair-wise crossings of those dimensions;
- to present these results in a tabular format for every test, along with normal averages and normal standard deviations for that test;
- to score several tests (for example, sentence production) which require non-numerical analysis;
- to facilitate and systematize human scoring along several dimensions which are difficult or impossible to formalize so that they can be scored automatically (for example, classification of errors in naming), and to store this scoring so that it need only be done once;

- to provide textual summaries of results on all relevant dimensions, noting especially results below norms;

- to highlight all notable associations and dissociations of results within a single test with textual commentary and, where appropriate, with theoretical explanation;

- to re-present selected data in tabular format in order to facilitate comparisons between two or more subtests;

- to highlight all notable associations and dissociations of results between two or more subtests with textual commentary and, where appropriate, with theoretical explanation.

The choice of a programming environment was determined entirely by the demands of the task. Many production system environments do not provide flexible tools for all computational tasks- for example, for manipulating free-format ASCII data files. Moreover, many tasks which can in theory be handled by productions can be more easily handled with standard procedural languages. In order to overcome these two limitations of production systems, it was necessary to find a production system which gave full access to an underlying programming language. I was constrained practically in having to work on the same platform as the computerized PAL battery, which ran only on an Apple Macintosh computer.

I spent several months analyzing different production system environments trying to find one which would allow such access. I was fortunate enough to discover Peter Shell and Jaime Carbonell's FRuleKit, a frame-based production system written in Common Lisp which is distributed with its source code. The system is somewhat similar to the 'classic' production system, OPS5, as FRulekit has a similar syntax and relies on an augmented version of the Rete pattern matcher that was developed for use in the OPS5 system (see Brownston et al, 1985). However, FRulekit is more flexible and more powerful than OPS5. FRulekit provides the inference engine which runs the production system in the analysis system.

The inference engine is used in a simple way, which is well suited to the linear task demands of the problem. The rules are stored in a modular fashion, allowing for rapid development and extension. A framework application which is built around the inference engine accesses those rule files in an ordered fashion, keeping the data flow through the production system easy to understand. The framework application looks for result files in a user-specified folder. If it finds that it has the data file or files which are to be analyzed by that rule set, then it loads the relevant rule file (see Figure 2.3). In this way, rules are not loaded unless needed. Once loaded, the rules in each set are fired (by forward chaining- i.e. matching on the left hand side only) in a linear fashion. This is a somewhat perverse limitation on the normal use of the rule-based system, which is to match rules dynamically as circumstances warrant. However, it has the advantage of allowing for easy control of the textual output, which is in this case is of paramount importance. Very few of the rules have any other purpose than to print a textual output.

FRuleKit does not provide any interface tools. I wrote a very simple interface to allow it to query users, which allows different kinds of questions to be asked, and allows for simple range-checks on the responses.

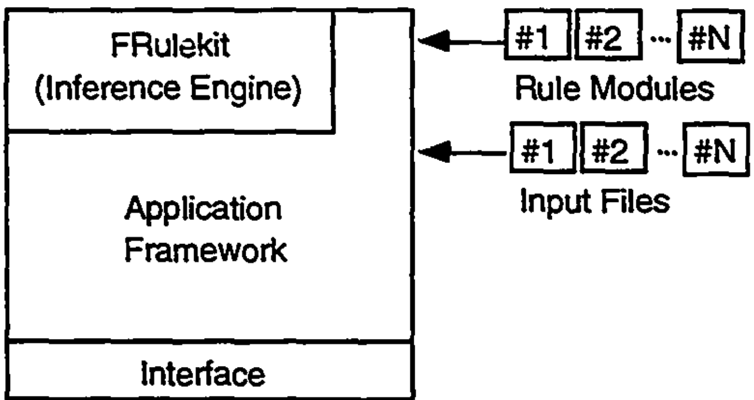


Figure 2.3
Diagram Of The Aphasia Battery Analysis System

The main application framework has three main components, each devoted to a different aspect of the analysis task. The first and simplest component gathers some basic information about the patient and the location of the patient's data files. The second component scores those few test results which need to be scored by human beings: the error results from the tests of reading, repetition, and both oral and written naming. All of these must be coded in a way which relies upon semantic knowledge, and which is thus impossible to automate fully (with the exception that errors of omission may of course be scored automatically). The application facilitates the process by presenting each erroneous result from each of these four tests, along with the stimuli which elicited it, and asks the user to choose from a list of possible codes that describe the patient's response. The frames which store the errors are written immediately to disk and automatically reloaded if the system is run again with the same data. The third component of the application is the analysis system proper, which analyzes each of the tests in turn. It first presents a table showing the raw results broken down by every level in the test, and by all pair-wise crossings of those levels. Any results which are more than two standard deviations below normal performance are flagged. The application then summarizes the results in paragraphs which group together relevant results. Finally, it loads the rule-file associated with that test, and fires rules by checking them for matches in a linear fashion. Any text which is returned by a matching rule is written to the standard output, most often the screen. An edited example of typical output from the system is included as Appendix A.

This use of a hybrid application/production system gave us access the two main benefits of rule-based systems: modularity of knowledge, and ease of expansion of the knowledge base. However, because the system could be extended in Common Lisp, it was also possible to tailor the system to our own needs, to perform many computational tasks which might be impossible or inefficient in a production system which was 'sealed' from a more flexible programming language, and to build in a number of auxiliary tools which have proven to be useful, for moving, renaming, re-formatting files, and for analyzing group performance on multiple result files.

iii.) Janus: A tool for knowledge engineering

The most difficult part of constructing the PAL Battery Analysis System was developing the rule base. It has long been recognized in artificial intelligence circles that the mapping between a 'chunk' of knowledge and a production rule or set of production rules is often surprisingly complex. As a result, it is often very difficult to translate an expert's real-world knowledge into the highly structured knowledge required by a production system's rule base. Unweaving the complexity of the relation between knowledge and rules can often be incredibly frustrating for both the expert and the knowledge engineer, as it was in our case. In order to address this problem, I wrote a knowledge engineering tool which I named 'Janus', after the Greek God who looks in two directions at the same time. The name captures the purpose of the tool, which was to look simultaneously into the 'computational world' of Lisp and the 'neurolinguistic world' of the PAL battery. Although the tool was quite simple, it turned out to be remarkably effective, speeding up the process of rule writing by at least an order of magnitude. In this subsection the tool is briefly described.

Although Janus does translate between a kind of pseudo-English and Lisp, its main purpose is not to translate the rules but rather to facilitate the structuring of the rule-base in hierarchical fashion which allows easy consideration of all permutations of a set of arguments. The program facilitates such structuring by forcing the expert (who may interact with it directly, without the need of an intermediary knowledge engineer) to focus upon a small set of arguments at one time, and by making it easy and natural to change those arguments in a systematic way.

The program's structure is diagrammed in Figure 2.4. The structure consists of two main parts: a series of 'data funnels', which narrow the focus to a manageable number of arguments, and a set of tools for writing and 'morphing' (systematically altering) rules which use the chosen set of arguments.

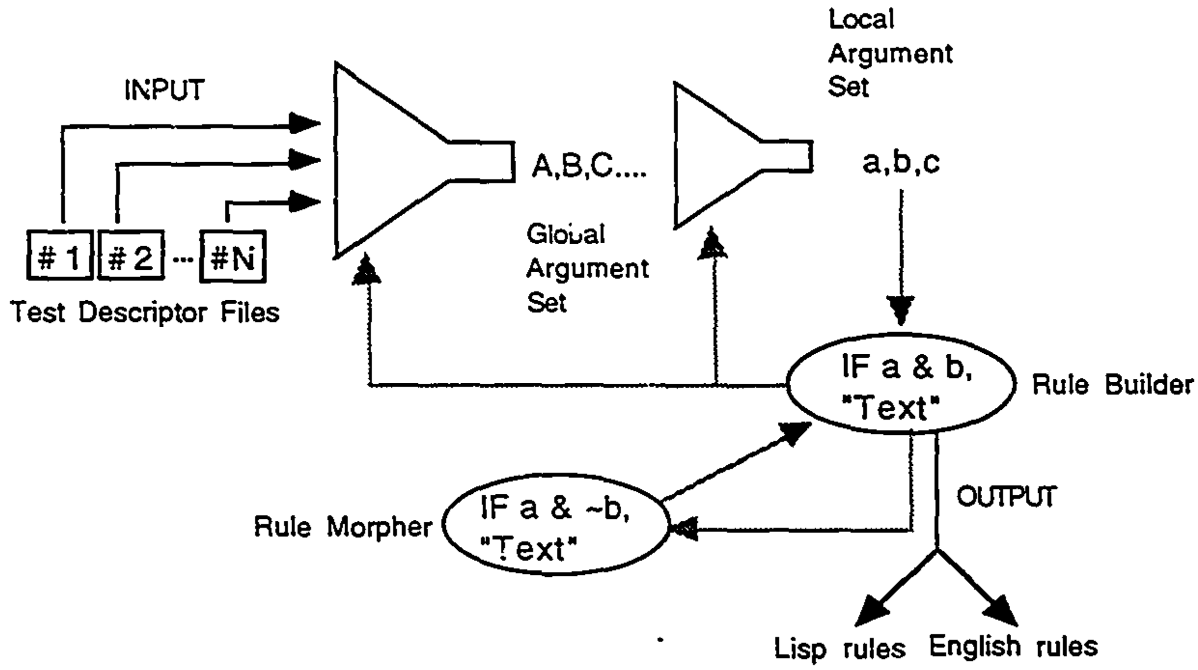


Figure 2.4:
Diagrammatic Representation Of Janus: A Knowledge-Engineering Tool
Black lines represent input and output routes.
Grey lines represent data flow.

To begin using the tool, a user selects one or more 'test descriptor files'. These are simply ASCII files containing a list of all possible arguments that a rule may use. In our case, each one contained a list of all markers contained in a single PAL test. Since some PAL tests contain many markers, this global argument set may be very large. Because the interface uses a list-based interface, it is usually convenient (but not necessary) to narrow down the focus by choosing a subset of the global argument set, the local argument set. The user may then choose whether he wishes to work with the norms, percent scores, exact scores or all three scores of that set. The local argument set is marked accordingly. With the resultant arguments, the user can build standard production system rules using the tools which have been built in for that purpose. These tools allow the user to build the conditions which comprise the left-hand side of rules. The rules are written in infix notation, which places the operator between the arguments as in standard English usage. Any of the arguments may be compared to any other argument, to built-in (and user-definable) constants, or to some mathematical combination of the arguments and constants. For example, a user may build the line following line:

```
WNaming:HF% < WNaming:HFNorm
```

This condition checks to see if the subject's percent score for writing names of stimuli with high frequency names was less than the normal average percent score for writing names of stimuli with high frequency names. Each of the items on either side of the mathematical sign can be chosen as a complete unit from the argument set defined within Janus.

As a second example consider the following line:

```
(Reading:HF% - Reading:LF%) >= 15
```

This condition checks to see if the subject's percent score for reading high frequency words was at least 15% higher than his percent score for reading low frequency words.

The rule builder contains a loop which allows the user to build as many conditions per rule as he likes in this fashion

When he has finished building the left hand side of the rule (the conditions), then the user selects an option which allows him to write the right hand side (the conclusion). Janus only allows the user to execute a single action on the RHS if the rule fires: to write a text to the screen. That text may contain any of the arguments of the rule, and may also make references to the subject's name (using a defined pseudonym) or sex, since such references are automatically translated into variables (whose value is fixed at runtime) when the rule is translated into the prefix notation of Lisp.

When the right hand side is written, the rule is ready to be translated. If the user accepts the rule as it appears, it is written out to two files, one containing the rule as it was written, and one containing the Lisp translation of that rule. An example of a rule written in each syntax is provided in Figure 2.5.

At any time a rule may be 'morphed' using a variety of built-in tools. Most often this is done after a complete rule has been written to disk. The tools allow the user to easily make a number of common changes to the condition on the left-hand side of the rule: to change the logical or mathematical sign in one or more of the conditions, to change one or more of the arguments which appear in the conditions, to alter or delete one or more of the conditions, to add one or more conditions to a stored database of frequently used conditions, to retrieve one or more conditions from that database, or simply to add a new condition to the rule. The user may also edit the right-hand side text, or start again from any point in the program: either reading in new arguments from disk, choosing a new subset of arguments from the current global argument set, or beginning anew with the current local argument set.

Since many rules in most production systems are small variations on other rules which have already been written, the rule-morphing capability at the heart of this program proved itself to be extremely useful. Hierarchical thinking is enforced in a painless and transparent manner, simply by encouraging the user to focus his attention on making incremental changes to the current rule before proceeding to a new rule. As the users, Bub and I deemed Janus to be an unqualified success. Our experience in using it convinced us that computer-aided knowledge engineering is well worth the effort needed to design the proper tools. In my concluding chapter I will make some suggestions about how this tool might be usefully extended.

```
RULE ACCESS-DISORDER-CROSS8
; CFW/BUB - 14:24:15 22/1/1992
IF
 AUDITORY-WP-MATCH.ALL% > 85
 AUDITORY-LD.HI% < 65
THEN
 Fnord does quite well in a test of word
comprehension using concrete words. However,
he fails to identify many concrete words in a
lexical decision task, incorrectly rejecting these
as non-words. We must conclude that the lexical
decision task in this case is not providing a clear
picture of his ability to map auditory input onto
word-forms. Fnord must still be carrying out
this procedure but he cannot reliably determine
the status of the word, presumably due to a
problem in fulfilling the requirements of a lexical
decision task.
```

```
(RULE ACCESS-DISORDER-CROSS8
; CFW/BUB - 14:24:46 22/1/1992
 :LHS
 ((AUDITORY-WP-MATCH [ALL PERCENT]
                         =PERCENT242)
 (AUDITORY-LD [HI PERCENT]
                 =PERCENT243
 (CHECK
 (AND ((> =PERCENT242 85)
       (< =PERCENT243 65))))))
 :RHS
 ((format t "~%~5t~a does quite well in a
test of word comprehension using concrete
words. However, ~a fails to identify many
concrete words in a lexical decision task,
incorrectly rejecting these as non-words.
We must conclude that the lexical decision
task in this case is not providing a clear
picture of ~a ability to map auditory input
onto word-forms. ~a must still be carrying
out this procedure but ~a cannot reliably
determine the status of the word,
presumably due to a problem in fulfilling the
requirements of a lexical decision task. "
 *NAME*
 *PRONOUN*
 *POSSESSIVE*
 *NAME*
 *PRONOUN*
)))
```

Figure 2.5:
An example of a rule as written by Janus (left hand side) and the same rule translated by Janus into executable Lisp (right hand side). The purpose of the rule shown is to test for the combination of being poor at lexical decision, and good at auditory word-picture matching of concrete words, an unusual dissociation which would cast doubt upon the lexical decision result.

iv.) Aphasia diagrams (A-grams)

Although the PAL analysis system greatly simplified the process of scoring and organizing the data collected from the 24 tests of the battery, it did not allow us to fully overcome a problem which plagues modern aphasiological research: the difficulty of meaningfully integrating all the data that must be collected from a single patient in a theoretically-complete aphasiological investigation. This difficulty arises mainly from a simple limitation of the human mind, which is our inability to hold many different factors in working memory simultaneously. A typical output file from the

PAL analysis system runs to over forty pages of text. Even though noteworthy dissociations are detected and flagged by the program, a human user still finds it very difficult to keep track of them all. I explored several methods of compact data representation in an effort to overcome this problem, and finally invented a representation which I dubbed an aphasia diagram, or an A-gram for short.

A-grams not only allow us to represent the results from all normed subtests in our aphasiology battery on a single one-page form, but they also allow us to quickly and easily spot the cross-test associations and dissociations which are vital to our understanding of the patient profile. The diagrams achieve this purpose by representing such associations and dissociations visually. They thus compensate for one of the human brain's greatest natural weaknesses- its annoyingly small short-term storage capacity- by exploiting one of its great natural strengths: its ability to 'automatically' parse visual patterns and pick out similarities between such patterns.

An A-gram consists of 22 diagrams, one for every subtest in the PAL except the sentence production tests- although when that test is well-normed it will be added. Although the conceptual relation between the 22 diagrams which constitute a single A-gram is quite complex, each individual diagram is itself very simple. We will consider as an example the two test result diagrams representing the results from tests of auditory and written word-picture matching, which are reproduced in Figure 2.6 below.

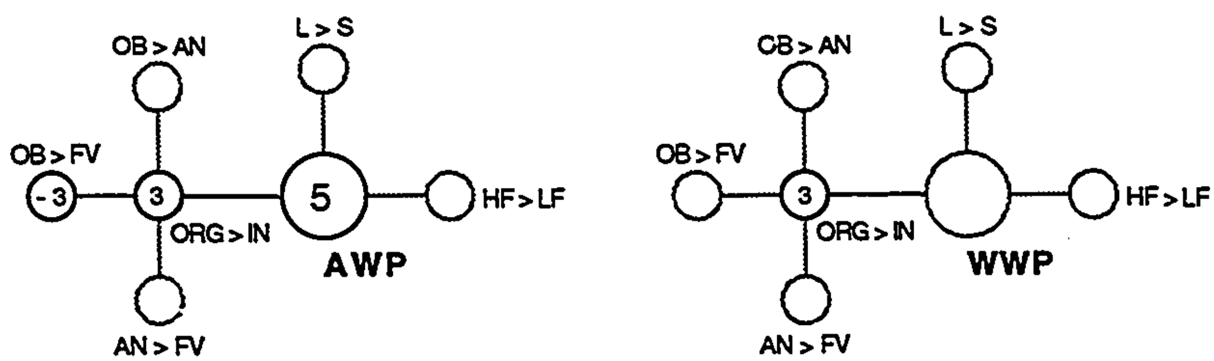


Figure 2.6:
Examples Of Two Test Result Diagrams
(22 such diagrams constitute a single A-gram)

Each test result is represented by a large central circle. If the patient scored above the two standard deviation cut-off point below the normal average score on the test which the circle represents, then the circle is presented in grey, and is empty. In the example above, the large circle on the right is grey, indicating that the patient scored within the normal range on a test of written word-picture matching. If the patient scored below the 2 SD cut-off point, then the circle representing the test is black, and contains a standardized score (the number of normal standard deviations below the normal average score on that test) which represents the patient's over-all performance on that test. By standardizing the scores we can compare results between tests with different statistical properties. In our example, the large circle on the left is black, and contains the number '5', indicating that the patient scored five standard scores below average on a test of auditory word picture-matching.

If the test was administered, but the patient was unable to complete it, then the circle representing that test would be black, and contain a dash instead of a number. If for any reason a test was never administered to a patient (or if the norms for the current version of the test render it unusable!) the circle is greyed out, but is filled with the letters 'N/A'.

Each circle may have zero or more arms radiating from it. Each arm represents a theoretically-relevant dissociation within the test. In the word-picture matching tests we are using as an illustration, there are three main dissociations of interest, thus there are three arms radiating from the main circle. The dissociations of interest are the effect of the frequency of the stimuli names, the effect of the length of the names, and effect of the category from which the stimuli are drawn (organic or inorganic). There are also three categorical subcomparisons of interest (objects versus animals, animals versus fruits and vegetables, and objects versus fruits and vegetables)- so the small circle at the end of the arm representing the organic/inorganic distinction itself has three arms radiating from it.

The numbers contained within the small circle represent a statistically significant dissociations (if any) and the direction of that dissociation. A significant dissociation in the direction of the label is indicated by a '3'. The symbol '3' is used simply because there are three types of dissociations which one may want to note, as explained below. In both tests in Figure 2.6, the patient was significantly better at matching names to pictures for organic than for inorganic stimuli. A significant difference in the direction opposite the label is indicated with a negative sign. For example, the '-3' in the 'OB > FV' (left) arm of the diagram representing the auditory word-picture test in Figure 2.6 tells us that the patient was significantly better at matching names to pictures for stimuli drawn from the 'fruits and vegetables' category than for stimuli drawn from the 'object' category.

As already mentioned, normal data needed to calculate the significance of some dissociations were not available. In such cases, I had to rely upon conservative ad hoc estimates of normal performance by assuming an average of 90% and a standard deviation of 10%. This limitation is indicated by representing dissociations with a '2' instead of a '3'. Cases in which the patient scored significantly low on one factor of a level, but within the normal range on the other (for example, if a patient were low on low frequency words but not high frequency words) were flagged by using a '1' in the circle at the end of the arm. These dissociations are reported for the sake of consistency with the literature, where such dissociations are commonly reported. However, it is clear that such dissociations are meaningless in themselves, since they can flag as significant a difference which is of negligible magnitude.

In all three cases directionality is flagged consistently, by appending a negative sign in front of the number if the directionality is in the direction opposite to the label.

If the dissociation was not significant, the entire arm is represented in grey, and the small circle at the end of the arm is left blank. One tentative conclusion is that the patient whose results are represented in Figure 2.6 has a category-specific sparing for the category of fruits and vegetables. We may infer this because the significant effect in favour of organic stimuli is due only to a significant advantage for matching the names of fruits and vegetables over matching the names of objects.

Such a tentative conclusion would of course need to be buttressed with parallel results from other related tests. The diagrams are designed to facilitate the making of the appropriate comparisons, because dissociations which are common to more than one test are always represented by arms in the same place, just as they are in the two examples in figure 2.6. Thus if there is also a category-specific effect of organicity in the auditory naming test, it will be represented by the identical black three-pronged left-facing arm that represents such a dissociation in the auditory-written picture matching test. In Figure 2.6, it is immediately obvious that the patient showed the same organicity effect (albeit without scoring low overall and without the fruit and vegetable advantage) on the written version as on the auditory version of the word-picture matching test.

Some tests cannot be interpreted without knowledge of the qualitative error pattern on that test. A-grams include relevant error information from seven error categories, which are represented by seven small labeled sets of squares on the far right of the form.

A-grams are simple in conception. They do not provide any new information or simplify the data-gathering phase. However, they do present information that has already been gathered in such a way as to render it easily comprehensible, allowing for quick and deep understanding of a single patient's test results, and greatly facilitating cross-patient comparisons. Despite their simplicity, they have turned out to be so useful to us that they have in effect supplanted much of the functionality of the PAL analysis system. I will have more to say about their utility, and about possible ways to increase it, in my concluding chapter.

EXPERIMENTAL METHODOLOGY

Subjects

Twenty-two subjects who were diagnosed with either probable PPA or probable DAT participated in this study. All eleven of the DAT and 10 of the eleven PPA patients were diagnosed by one of five referring neurologists who specialized in cognitive neurology, and were familiar with the literature on both DAT and PPA. Since the patients came from different sources, we were not able to specify the diagnostic criteria used. The final PPA patient (AB) was referred with a tentative diagnosis of PPA by a neurologist who did not specialize in cognitive neurology. In that case the diagnosis was confirmed by Bub and myself on the basis of both a neuropsychological assessment and the subject's performance on the PAL battery.

Subjects were screened informally with a simple test designed to ensure that they could hear the auditory stimuli and see the visual stimuli. No subjects were eliminated with this screening.

As noted above, it was not possible to match subjects on demographic variables due to the difficulty of finding sufficient appropriate subjects. The PPA group was significantly younger than the DAT group ($t(20) = -2.26$ $p < 0.05$). There was no significant difference in years of education ($t(17.1) = 1.81$ $p > 0.05$). The groups were about equally balanced on sex (PPA: Six males, five females; DAT: Five males, six females). There was a significant difference in years since symptom onset ($t(20) = 2.6$; $P < 0.05$). However, this measure is confounded with a selection bias in two ways, both because it is impossible to test DAT patients late in the disease process, and because it is generally not as easy to date the early symptoms of DAT as it is to date the early symptoms of PPA, since the former tend to be more subtle. Because of this, the estimate of years since onset for the DAT patients is more likely to represent an underestimate which is measured from the first medical diagnosis rather than from the initial appearance of symptoms.

		PPA					DAT		
NAME	AGE	SEX	EDU.	Years since onset	NAME	AGE	SEX	EDU.	Years since onset
AB	72	F	10	4	NB	86	M	11	2
DM	57	M	16	6	AB	81	F	14	2
JD	69	M	8	1	DD	69	M	18	2
CD	79	F	14	1/12	EF	76	F	7	2
JH	59	M	16	7	IK	60	M	15	2
BH	74	F	11	1	DD	78	M	5	2
BL	70	F	16	3	MR	63	F	12	2
JL	70	M	16	8	DS	79	F	11	2
CM	50	M	16	3	JS	80	M	13	2
ES	79	F	12	5	RS	86	F	16	1
MW	66	M	12	3	YS	82	F	4	1
Ave.	67.7		13.4	3.7	Ave.	76.4		11.5	1.8
SD	9.1		2.9	2.7	SD	8.7		4.5	0.4

Table 2.2
Basic demographic data for
the 22 patients in this study

Materials

Each of the tests in PAL was presented in a consistent manner, using PsychLab software running on either a Macintosh LC computer attached to a 14 inch Apple monitor, or, more often, a Powerbook 120 Macintosh portable computer. Patients who were tested in a hospital in Montreal were tested with the LC. Patients who were tested at their own homes or at a hospital in another city were tested using the portable computer. All but four of the patients were tested by the same tester. Those four were tested by two research assistants who been trained to administer the tests in a standardized fashion.

Stimuli size did not vary with screen size. All visual stimuli were presented in black and white. Auditory stimuli were recorded monophonically by a female speech therapist using 8 bit sound, and played to the subject using either a Aiwa SC-A9 or a Yamaha YMT-S10 speaker.

Procedure

Some of the patients were seen as part of a clinical consultation. Others participated voluntarily for research purposes. All patients were informed as to the purpose of the study and understood that they would be given a set of computerized tests that had been designed to assess their language functioning.

The tests were presented in a roughly standard order, although deviations from that order were sometimes required for practical reasons. Within most tests, the order of stimulus presentation was automatically randomized for each subject. The only exceptions were the sentence production tests, which were presented in a consistent order because they include two unscored 'practice items' which must be given first.

Responses were either given verbally, by pointing, or (where required) by writing, and were typed in to the computer by the tester. Since competence rather than performance was the focus of the testing, subjects were informed that they could ask for any test item to be repeated as many times as they liked. They were also allowed to change answers immediately after giving them.

Testing required from 2 to 10 sessions which varied in length from thirty minutes to about three hours. The total time spent testing a single patient varied between five and fifteen hours, due to wide variations in patient stamina and speed.

Scoring

In order to facilitate comparisons between tests, each test result in this thesis is presented (as in the A-grams) as a standard score, in terms of the number of normal standard deviations below the normal average score for that test. The rare negative scores in the tables or the A-grams therefore represent scores above the normal average. A score was considered significantly low if it was two or more standard deviations below the normal average (that is, if it was in the bottom 4% of the normal population).

Dissociations were also calculated in terms of standard scores. Scores from two levels of a single test were considered significantly different if their difference was more than two normal standard deviations larger than the average normal difference. Because some levels of a factor contain a small number of items, significant differences were sometimes obtainable using this formula as the result of a difference of a single item. Such differences were excluded from consideration, in order to help minimize Type I error. The reasoning is straightforward: by definition, Type I errors will be due to random error. Since random errors are (by definition) independent events, the chances of a single such error will be higher than the chance of more than one error. Therefore, there is a greater chance of Type I error due to a difference of single random error than to a difference of more than one random error. The reasoning might be extended: in general, the smaller the number of items upon which a significant effect is based, the greater the chance of a Type I error. The decision to ignore errors due to a single item is equivalent to choosing a more conservative alpha level (the acceptable probability of Type I error) in a specific identified situation in which Type I error is most likely to occur.

In a few cases, the stimuli set which was seen by the normals differed slightly from the set used in the study (as described in some detail in the first section of this chapter), resulting in some levels of variables for which normal performance is not documented. In these cases, a level of 'clinical significance' was defined. Scores from two levels of a single test were considered 'clinically significant' (in a rather ad hoc manner, which is, however, consistent with standard clinical practice) if they differed by more than 15%.

Methodological Caveats

There are three potential problems with this methodology which deserve special recognition. The three are related inasmuch as they represent different aspects of problems with statistical power: that is, with the management of random error.

The first problem is that there were very few patients in the study. While there are excellent practical reasons for limiting the number of patients in studies like this one, the low N renders any claims to have established stable subtypes with a single diagnostic group, or significant differences between the two groups to be of statistically low reliability.

The second problem is that there were a number of comparisons which were based upon a low number of observations. This problem is particularly acute when comparisons were made with normal performance. Because the normal scores were usually quite high, significant cut off scores may often reflect errors on only a few items. In order to minimize this problem, in this study all comparisons between factors with less than eight observations per cell have been ignored. I thus have almost nothing to say about two-way interactions. I have also (as mentioned above) ignored out any dissociations which were significant, but dependent upon a difference of only a single item. Nevertheless, the small number of observations in many cells remains a problem.

The third problem is more apparent than real: the problem of statistical power. About 130 comparisons per subject were examined in this study (see Chapter 6 for information on how this number was derived). Only the use of extremely conservative significance levels can justify this many comparisons per patient. In this thesis I generally relied upon a cut-off point of 2 SDs (the 96% confidence interval) for identifying significant effects. An alpha level of 4% will lead to approximately five 'randomly' significant effects per patient. Although it might be argued that this is not conservative enough given the number of comparisons, it was felt that the exploratory nature of the work did not necessitate extreme statistical conservatism. Furthermore, a post-hoc examination of the significance levels proved that the vast majority of the identified effects would have held up under much more conservative measures. Among the PPA patients, 86.8% of all effects which were significant at the 0.23% (2 SD) cut off point would also have been identified as significant at the 0.001% (3 SD) cut-off point. 77% would have been significant at 0.00003% (4 SD) level! Although the significant results among the DAT patients were not quite so robust, most effects would have been identified as significant at both the 3 SD cut off point (74.6%) and the 4 SD cut off point (61.3%). These post-hoc analyses do not totally dismiss the problem of Type I error, but they do suggest that the results reported in this thesis would not be highly sensitive to a more conservative alpha level.

When cognitive neuropsychology adopted the single patient method, it also adopted all three of these problems. The entire strength of statistical reasoning lies in its ability to factor out random effects which may affect results. When we can not factor out random effects statistically, we must rely on other tools to do so, for to fail to rule them out at all is to fail to do science. The tool for ruling out random effects in cognitive neuropsychology is logical inference. Just as we do not need to sample randomly from a set of incorrect logical syllogisms to prove their incorrectness statistically, so there is no need to rely upon statistical methods to argue that a set of 'impossible' deficits is impossible, so long as we have sufficient confidence in the model which is used to define the possible.

In this study, I relied for the purposes of clarifying my inferences upon the model of word processing which is implicitly instantiated in the design of the tests which constitute PAL, based on Howard and Franklin's (1988) adaptation of a model of single word processing. Howard and Franklin's model is reproduced in Figure 2.7. The model serves mainly as mnemonic of what is known to be possible, because it was developed as an attempt to capture in as much detail as possible what is known about general human competence in language production and comprehension. Each box and line represents a functional element or a connection between elements which has been shown to fail independently from the others. Most such failures correspond to known syndromes. Many syndromes which are conceivable are implicitly ruled out in practice by this model, since certain sets of deficits will have very non-parsimonious explanations (and some will have no explanation at all) under this model. Other observed syndromes are illuminated by the model's instantiation of knowledge about how identical symptoms can derive from different underlying causes (for an interesting discussion of this issue, see Coltheart et al, 1987). I will not justify or explain the model in detail here (see Morton, 1979, Monsell, 1987, and Howard & Franklin, 1988) but I will refer to it as necessary in discussing the individual patients.

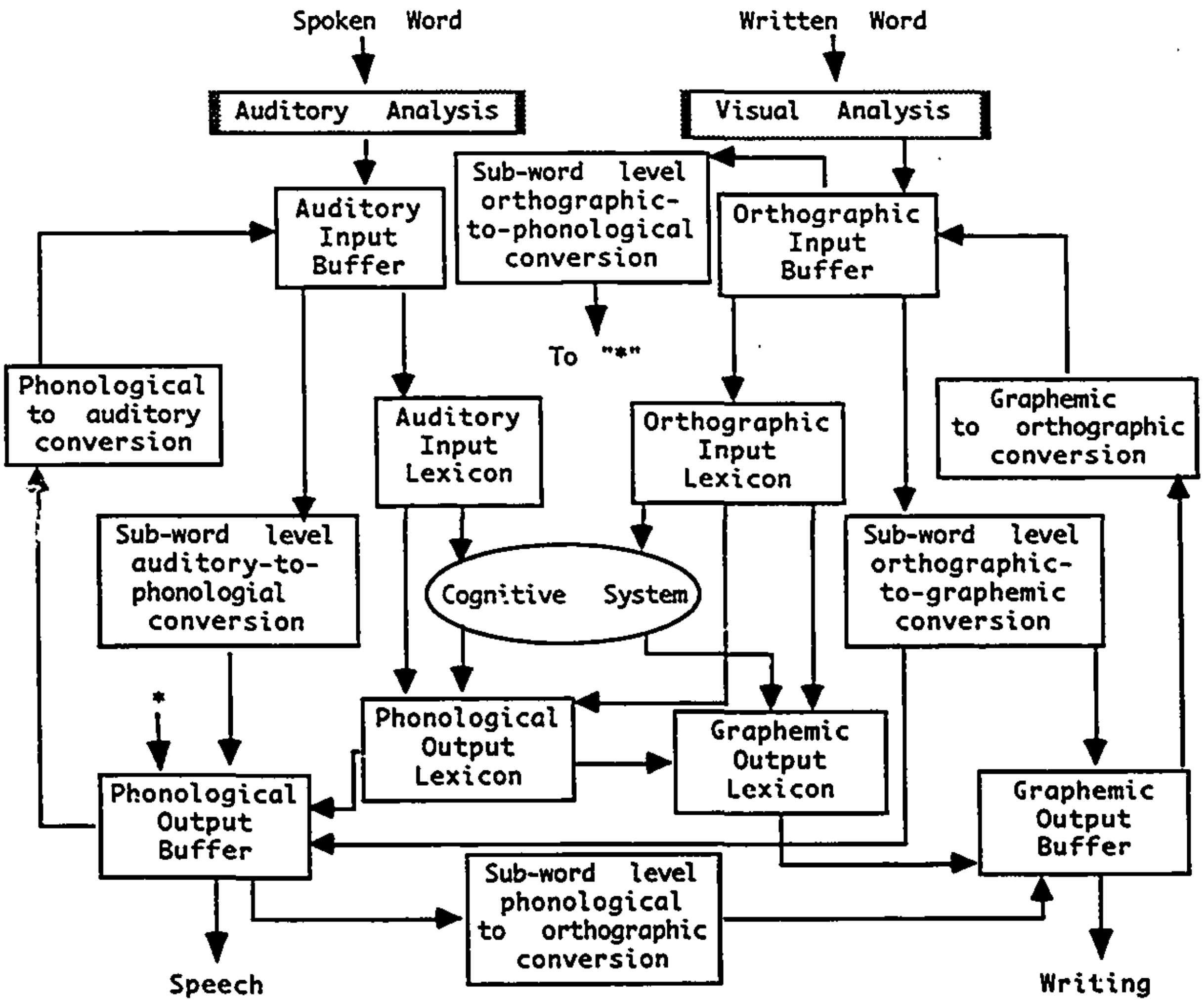


Figure 2.7
A model of single word processing
(adapted from from Howard and Franklin, 1988)

Using such a model to guide our thinking helps us move from the level of the merely descriptive, in which 'anything goes' and all observations must therefore be taken to be of equal importance, to the level of theory-driven description, in which we are able to focus our attention on findings which are most likely to lead to scientific insights (See Popper, 1959 and Kuhn, 1970. The specific role of models in cognitive neuropsychology is discussed in Shallice, 1988. The dangers inherent in such a model-based approach are perhaps best illustrated by its opponents- see Fort, 1941/1974 for a genuinely illuminating, if unorthodox, historical analysis of the limitations of the model-based approach, and Feyerabend, 1988, for a more formal argument about the dangers of the approach.). Although such model-guided reasoning does not totally overcome the problems caused by the lack of statistical power, it greatly limits their danger.

I have kept this discussion of methodological issues in cognitive neuropsychology brief, because the subject has been considered in subtle detail elsewhere. I refer the interested reader particularly to Chapters 9 and 10 of Shallice (1988).

Chapter Three:
Primary Progressive Aphasia: Review

Not chaos-like together wash'd and bruis'd,
But, as the world, harmoniously confus'd:
Where order in variety we see
And where, though all things differ, all agree.

Alexander Pope
Windsor Forest

In this chapter, the current understanding of PPA is explored in some depth. I reviewed data from 170 published contacts with 112 patients, in order to document the general clinical, biographical, neuroanatomical, and neuropsychological profile of patients with the disorder, and to make a rough attempt to chart its progression as it affects the five most-frequently examined general linguistic skills: oral and written naming, reading, repetition, and general comprehension. My review suggests that these five skills are differentially affected by the progression of the disorder.

METHOD

I reviewed every English and French language paper which purported to concern a primary progressive aphasic, excluding papers which referred to patients who were seen before Mesulam defined the syndrome. I included papers on semantic dementia, a subtype of progressive aphasia which spares syntactic abilities while affecting naming, word comprehension, and reading (see Hodges et al, 1992), since very few progressive aphasics have had the semantic testing necessary to differentiate them from patients with semantic dementia. Papers describing patients who were later shown to have a known progressive disorder other than PPA were not excluded, since I was interested in progressive aphasia as a syndrome, and since the known rate of misdiagnosis (as documented below) suggests that many PPA patients who have not yet been shown to have a known progressive disorder do in fact have one.

In order to allow for the progressive nature of the disorder to be examined, the neuropsychological data were encoded in the following way. First, each test result reported was mapped to a single digit code from 0 to 3. A code of 0 was given if the test result was reported to be within the normal range, or if the test result were within 1 standard deviation of the average when standard deviations were reported. A code of 1 was given if the patient was described as 'mildly impaired', or if the reported test scores fell below one but above two standard deviations below normal scores. A code of 2 was assigned if the patient attained a score described as 'moderately impaired', or if his score fell below two but above three standard deviations from the norm. A code of 3 was assigned if the patient's performance on a test was described as 'severely impaired', or if his scores were more than three standard deviations below normal scores.

In cases in which the patient's performance was reported to be impaired but no scores were provided, a code of 1 was conservatively assigned. Codes drawn from numerical data took precedence over codes assigned from textual descriptions.

Data was coded by patient, since there are several patients who have been reported in the literature more than once.

It proved impossible to code data from many papers using the above criteria, since papers were occasionally published by authors who neglected to report standard scores or even assign textual severity labels to patient performance. In cases where no code could be assigned using the above criteria, an assumption was made that the test used had a normal mean of 90% and a standard deviation of 10%. The results were then coded accordingly, as described above. Although these assumptions about normal performance are clearly ad hoc, they have some heuristic value. The assumed mean and standard deviation are extremely conservative for the kinds of tests reported, since normal performance on these tests is usually near ceiling. Any error introduced by the assumptions can thus be reasonably supposed to fall on the side of underestimating symptom severity.

Although the raw encoded data and the raw average coded scores are reported in this paper, a second set of results for each test is also reported, in which a rule of 'retroactive normality' was applied. This rule made the assumption that if a patient scored in the normal range on a test at a given reporting period, then that same patient would also have scored in the normal range if he or she had been given the same test at an earlier period. In applying the rule, every coded score of zero which was found at least one year after symptom onset was 'carried back' to every previous year since that patient's symptom onset. The application of this rule simply helps to clarify the real pattern of degeneration within the entire population, compensating somewhat for the skew towards pathology which is due to the fact that only results documenting deficits are publishable.

The data were encoded into eleven reporting periods, one for each year since symptom onset from 0 to 9, and one for all test results obtained more than 10 years after onset. When more than one result was reported within a single test period, the most severe result was used.

A twelfth category was used to code all results which could not be related to the time since symptom onset.

Results for all reported tests were coded in five broad classes: tests of oral naming, written naming, reading, repetition, and comprehension. Results from tests of comprehension are particularly difficult to compare, as opinions about what constitutes a test of comprehension vary more widely than opinions about testing the other abilities. For the purposes of this paper, any result of a test claimed as a test of comprehension in either the written or spoken modality was encoded as a test of comprehension.

RESULTS

Demographic Data

Data was reported for 112 patients: 73 (66%) males and 39 (34%) females. Of the 70 (62%) for whom handedness was specified, 68 (97%) were right-handed, and 2 (3%) were left-handed. 80 (71%) patients had only a single contact reported, with the remainder (29%) ranging up to 6 contacts over an average of 3.3 years from the initial contact. The average time since symptom onset for those patients who for whom only a single contact was reported was 4.3 years.

	AVERAGE	SD	HI	LO	N
MALE	-	-	-	-	73
FEMALE	-	-	-	-	39
RIGHT-HANDED	-	-	-	-	68
LEFT-HANDED	-	-	-	-	2
HANDEDNESS UNSPECIFIED	-	-	-	-	43
NUMBER OF CONTACTS	1.5	1	6	1	170
TOTAL SYMPTOM SPAN	5	2.9	15	1	81
SPAN MEDICALLY FOLLOWED	3.3	2.6	11	1	33
SINGLE-CONTACT SPAN	4.3	2.8	14	1	80
AGE WHEN SEEN	63.4	9.4	80	17	97
TIME SINCE ONSET WHEN SEEN	3.3	2.8	14	0	112
AGE AT ONSET	59.1	9.6	80	17	82
MALE	59.3	7.7	77	40	50
FEMALE	58.9	12	78	17	32

Table 3.1: Demographic data of all patients
Number of contacts = Number of data collection contacts reported
Total symptom span = Years from symptom onset to last reported contact
Span medically followed = Years from first contact to last reported contact (if seen more than once)
Single-contact span = Years since onset for patients with only one contact reported.

In a 1987 description of his original six patients, Mesulam wrote that in the terminal stages of PPA, cognitive impairments and other signs of dementia might be present, but "not before eight to twelve years after onset" (Mesulam, 1987, pp.553), implying that a necessary component of the syndrome is an initial period of at least eight years during which language-related deterioration is the only mental abnormality. Mesulam's suggestion does not appear to be widely accepted. Only 17 (15%) of the 112 patients for whom time since symptom onset is reported were ever seen eight or more years after symptom onset. Weintraub et al (1990) and Kertesz (1994) have both suggested that a period of two years of purely language-related deterioration is sufficient making the diagnosis. Seventy-six (68%) of the patients whose span since symptom onset is reported met this criteria.

The average age of symptom onset for all patients was 59.1 years (SD = 9.6 years). The modal value is 64 years. The average age of symptom onset did not differ significantly for males and females (p(1,80) = 0.884). The great majority of the patients (98, or 87%) reported were native English speakers. The rest comprised eight (7%) French speakers, three (3%) German speakers, and one Japanese, one Dutch, and one Italian speaker. The demographic data are presented in more detail in Table 3.1.

Presenting problem

The most commonly reported presenting problem was word-finding difficulty, which was reported by 42 patients (59% of the patients for whom the presenting problem was mentioned, and 37% of the total population). The next most common problem was a comprehension deficit, which was mentioned 22 times, accounting for 32% of the patients for whom the presenting problem is known, and 20% of the total population. Naming deficits, mentioned 16 times, were a problem at presentation for 14% of the total population, and thus for 23% of the reporting population. Other problems mentioned at presentation included speech hesitancy, phonemic paraphasias, dysarthria, slowed speech, and stuttering. These presenting problem data are summarized in Table 3.2.

PRESENTING PROBLEM	N	%	R%
UNSPECIFIED	42	37	59
WORD-FINDING	31	27	44
COMPREHENSION	22	20	32
NAMING	16	14	23
PEOPLE	6	5.3	8.5
OBJECTS	4	3.5	5.6
PLACES	2	1.8	2.8
UNSPECIFIED SPEECH	12	12	18
WRITING	6	5.3	8.5
HESITANCY	4	3.5	5.6
ARTICULATION	4	3.5	5.6
PHONEMIC PARAPHASIA	3	2.7	4.2
READING	3	2.7	4.2
PRONUNCIATION	2	1.8	2.8
NONFLUENCY	2	1.8	2.8
SLOW EFFORTFUL SPEECH	2	1.8	2.8
STUTTERING	1	0.9	1.4
SEMANTIC PARAPHASIA	1	0.9	1.4

Table 3.2: Presenting problems
% = Percent of total population
R% = Percent of those patients for whom the presenting problem is known

Neuroanatomical data

Data related to the site and nature of the brain degeneration was compiled from three sources: reports of MRI scans, reports of PET and SPECT scans, and reports from autopsy studies.

MRI

A total of 104 MRI scans were reported, of which 87 (84%) revealed a structural anomaly. These anomalies were classified by site for each of the twelve time measurement periods. The scans are summarized by hemisphere and time period in Table 3.3. 49 (56%) of the 87 positive studies reported anomalies in the left hemisphere only. The remaining 38 studies (43%) reported bilateral anomalies.

Years Post-Onset	0	1	2	3	4	5	6	7	8	9	9+	?	TOTALS
Total	4	7	15	11	18	13	5	2	4	2	4	20	105
Left Only	2	6	7	5	7	4	3	0	2	1	1	11	49
Left & Right	1	1	6	5	8	6	2	2	1	1	2	3	38
Negative	1	0	2	1	2	3	0	0	1	0	1	6	17
Positive	3	7	13	10	15	10	5	2	3	2	3	14	87

Table 3.3: MRI Anomalies
Overall Summary

Comparisons between studies are made difficult by the lack of a common vocabulary for reporting results. Thus, for example, one study might report 'a widening of the Sylvian Fissure' where another might have reported the identical anomaly as 'a tempo-parietal anomaly' or 'evidence of degeneration in the superior temporal lobe'. I did not make an attempt to unify the vocabulary, and report the results here in the terminology of the original descriptions.

The 50 scans which found exclusively left hemisphere anomalies are summarized in Table 3.4. The total number of measurements in this table does not sum to the total number of scans, since a single scan often uncovers anomalies at more than one site.

Years Post-Onset	0	1	2	3	4	5	6	7	8	9	9+	?	TOTALS
Sylvian Fissure	1	3	3	2	6	3	4	0	2	1	1	4	30
Ventricies	0	1	3	1	3	0	0	0	0	0	0	0	8
Lateral	0	1	1	1	2	0	0	0	0	0	1	0	6
Fronto-temporal	0	0	0	1	0	0	0	0	0	0	0	1	2
Entire hemisphere	0	1	0	0	0	0	0	0	0	0	0	0	1
Temporo-parietal	0	0	0	0	1	0	0	0	0	0	0	0	1
All three	0	0	0	0	0	0	0	0	0	0	0	1	1
Temporal lobe	1	3	5	3	1	1	0	0	1	0	1	4	20
Pole	0	0	0	2	0	0	0	0	1	0	0	0	3
Anterior	0	0	1	0	1	1	0	0	0	0	0	1	4
Superior	0	0	0	1	0	0	0	0	0	0	0	0	1
Frontal lobe	0	0	1	1	0	3	1	0	1	1	0	1	9
Frontal Horn	0	0	0	1	0	0	0	0	1	1	0	0	3
Pars Triangularis	0	0	0	0	0	3	0	0	0	0	0	2	5
Frontal fissure	0	0	0	0	0	0	1	0	0	0	0	0	1
Premotor area	0	0	0	0	0	0	0	0	0	0	0	1	1
Parietal Lobe	0	0	1	0	0	0	0	0	0	0	0	0	1
Subcortical	0	0	1	0	0	0	0	0	0	0	0	0	1

Table 3.4: MRI Anomalies
Left Hemisphere only, by site,
by years from symptom onset

The most commonly reported sites of anomaly among the MRI scans which found evidence of exclusively left hemisphere degeneration were the Sylvian Fissure (N = 31) and the temporal lobe (N= 20). Evidence of anomaly in the frontal lobes was reported less than half as often (N = 9), while evidence of a parietal lobe anomaly was mentioned only once.

Years Post-Onset	0	1	2	3	4	5	6	7	8	9	9+	?	TOTALS
Unspecified	1	0	2	4	1	3	1	1	1	1	2	1	18
Temporal lobe	0	1	0	0	1	2	0	1	0	0	0	1	6
Pole	0	0	0	0	1	0	0	0	0	0	0	0	1
Anterior	0	1	0	0	0	1	0	0	0	0	0	0	2
Inferior	0	0	0	0	0	0	0	0	0	0	0	0	0
Sylvian Fissure	0	0	0	2	2	0	0	0	0	0	0	0	4
Cerebellum	0	0	0	1	0	0	0	0	0	0	0	0	1
Ventricles	0	0	0	0	1	0	0	0	0	0	0	0	1
Subcortical	0	0	1	0	0	0	0	0	0	0	0	0	1
Frontal lobe	0	0	2	1	2	0	1	0	0	0	0	0	6

Table 3.5: MRI Anomalies
Bilateral, by site,
by years from symptom onset

Descriptions of the 38 MRI scans which found bilateral degeneration (Table 3.5) also mentioned the temporal lobe (N = 6) and the Sylvian fissure (N = 4) as the most common site of degeneration. The frontal lobes were mentioned just as often (N = 6). However, 18 (47%) of the scans mentioned bilateral degeneration without specifying the location.

PET & SPECT

Data from 59 PET and SPECT scan studies of PPA patients were reported. Fifty-seven (97%) of these studies reported finding anomalies in blood flow. These anomalies were classified by site for each of the twelve time measurement periods. The 59 studies are summarized by hemisphere and time period in Table 3.6.

Years Post-Onset	0	1	2	3	4	5	6	7	8	9	9+	?	TOTALS
Total	1	6	6	8	12	6	3	3	1	1	2	10	59
Left Only	1	4	3	6	9	3	3	0	1	0	1	8	39
Left & Right	0	2	3	1	3	3	0	3	0	1	1	1	18
Negative	0	0	0	1	0	0	0	0	0	0	0	1	2
Positive	1	6	6	7	12	6	3	3	1	1	2	9	57

Table 3.6: PET- and SPECT-verified anomalies
Overall summary

Forty (69%) of the 58 scans which reported abnormal results showed evidence of anomalies in the left hemisphere only. The remaining 18 studies (31%) showed evidence of bilateral anomalies. Note that the total number of measurements in each of these tables does not necessarily sum to the total number of scans, since a single scan can uncover anomalous findings at more than one site.

Years Post-Onset	0	1	2	3	4	5	6	7	8	9	9+	?	TOTALS
Fronto-temporal	0	1	0	0	1	0	0	0	0	0	0	3	5
Fronto-parietal	1	0	0	0	0	0	0	0	0	0	0	1	2
Temporo-parietal	0	0	0	1	1	0	2	0	0	0	0	2	6
Temporo-occipital	0	0	0	0	1	0	0	0	0	0	0	0	1
Perisylvian	0	0	0	0	1	0	1	0	0	0	0	0	2
Temporal lobe	0	0	2	4	3	6	2	0	0	1	1	1	20
Anterior	0	0	1	0	1	0	0	0	0	0	0	0	2
Posterior	0	0	0	0	0	0	0	0	0	0	1	0	1
Superior	0	0	1	0	0	1	0	0	0	0	0	0	2
Parietal Lobe	0	1	0	2	1	1	0	0	1	1	1	0	8
Anterior	0	0	0	0	1	0	0	0	0	0	0	0	1
Inferior	0	0	0	0	0	1	0	0	0	0	0	0	1
Frontal lobe	0	1	2	4	3	3	1	1	1	1	1	0	18
Posterior	0	1	1	0	1	0	0	0	0	0	0	0	3
Inferior	0	0	0	1	0	1	0	0	0	0	1	0	3
Occipital	0	0	0	0	0	0	0	0	1	0	0	0	1
Subcortical	0	0	0	1	0	3	1	0	0	0	0	0	5
Thalamus	0	0	0	0	0	2	1	0	0	0	0	0	3
Basal Ganglia	0	0	0	0	0	1	0	0	0	0	0	0	1
Insula	0	0	0	0	0	1	0	0	0	0	0	0	1
Caudate	0	0	0	0	0	1	0	0	0	0	0	0	1

Table 3.7: PET- and SPECT-verified anomalies
Left Hemisphere only, by site,
by years from symptom onset

Twenty-three (58%) of the 40 studies reporting unilaterally left results (Table 3.7) found anomalies in the left temporal lobe. In addition, another 12 (30%) reported anomalies in one of the fronto-temporal region, the temporo-parietal region, or the temporo-occipital region, and two (5%) specified perisylvian region involvement. Eighteen (45%) of the studies reported frontal lobe involvement, with an additional seven (18%) reporting either fronto-temporal involvement or fronto-parietal involvement. Only eight (20%) of the studies reported evidence of parietal lobe anomalies, though an additional eight (20%) reported evidence of either fronto-parietal or temporo-parietal involvement. A single study (3%) reported evidence of anomalies in the left occipital lobe.

Years Post-Onset	0	1	2	3	4	5	6	7	8	9	9+	TOTALS
Frontal & Temporal	0	2	2	1	3	1	0	0	0	0	0	9
All three	0	1	0	3	1	1	1	0	1	0	1	9
Temporal Only	0	1	1	1	4	0	0	0	0	0	0	7
Frontal Only	0	0	0	1	1	1	0	1	0	0	0	4
Frontal & Parietal	1	0	0	0	0	0	1	0	0	0	0	2
Temporal & Parietal	0	0	0	0	1	0	1	0	0	0	0	2
Total	1	4	3	6	10	3	3	1	1	0	1	33

Table 3.8: PET- and SPECT-verified ancmalies
Left Hemisphere only, by scan,
by years from symptom onset

In Table 3.8, the 33 scans which discovered unilaterally left-sided anomalies are presented by patient, in order to allow better understanding of the pattern of distribution of the anomalous findings.

No study reported finding exclusively parietal or occipital involvement. In every case in which involvement of either the parietal or the occipital lobe was reported, there was also temporal or frontal involvement. Two cases were reported (6%) in which both temporal and parietal lobe anomalies were found without any frontal lobe anomalies, and two cases (6%) in which both frontal and parietal involvement were found without temporal involvement. Nine studies (27%) reported frontal and temporal involvement without parietal involvement. Nine other studies (27%) uncovered evidence of involvement of the frontal, temporal, and parietal lobes.

Years Post-Onset	0	1	2	3	4	5	6	7	8	9	9+	?	TOTALS
Temporal lobe	0	2	0	0	0	0	0	1	0	1	0	0	4
Inferior	0	1	0	0	0	0	0	0	0	0	0	0	1
Parietal Lobe	0	0	1	0	0	1	0	1	0	0	0	0	3
Frontal lobe	0	1	2	1	2	3	0	2	0	0	1	1	13
Anterior	0	0	1	0	0	0	0	0	0	0	0	0	1
Inferior	0	1	0	0	0	1	0	0	0	0	.0	0	2
Pre-central	0	0	0	0	1	0	0	0	0	0	0	0	1

Table 3.9: PET- and SPECT-verified anomalies
Bilateral, by site,
by years from symptom onset

The break-down of information from the 18 scans with bilateral anomalies (Table 3.9) shows a different pattern. The most common reported site of bilateral anomaly was the frontal lobe, which was reported in 13 (72%) of the studies. Bilateral anomalies of the temporal lobe were reported in only four (22%) of the studies, and bilateral parietal lobe anomalies were reported in only three (17%) of the studies.

Autopsy Data

Autopsy data was available for 16 patients. The average time since symptom onset at death was 5.5 years (SD = 2.2). Three (19%) of the patients were deemed on the basis of autopsy results to have suffered from DAT; two (13%) patients were judged to have had Pick's Disease, and one (6%) was diagnosed as having had Creutzfeldt-Jakob disease. Of the remaining ten patients, six (38% of all autopsied patients, and 60% of all autopsied patients not given another diagnosis) showed 'spongioform changes' in layers II and III. Four of these patients also had neuronal loss, mainly limited to the frontal and temporal lobes, and the fronto-parietal area. The remaining four patients (25% of all autopsied patients, and 40% of all autopsied patients not given another diagnosis) did not have spongioform changes. Relatively little information is available for these patients. The autopsy data for the ten patients who did not receive a diagnosis other than PPA are summarized in Table 3.10.

SOURCE					
DURATION	8	5	8	10	3
ATROPHY	Lateral & 3rd ventricle dilated bilaterally	Slight atrophy of frontal lobes	Mild diffuse cortical atrophy, especially r. & l. frontal & sup. temporal gyri. Spared pre- & post-central gyri.	2 focal areas in l. inf. frontal & sup. temporal gyri.	Symmetric cerebral & cerebellar atrophy.
NEURONAL LOSS		Mild		Limited mainly to layer II	
PLAQUES	Yes	No	None		No
TANGLES	Yes	No	Rare		No
SPONGIOSIS	In nucleus basalis of Meynert, l. amygdala, l. sup. temporal lobe	Spongiform change in layer II			l. frontal lobe, with changes both intra-cellular & extra-cellular
PICK BODIES		No			
ASTROCYTOSIS / GLIOSIS	Hypertrophic astrocytes in nucleus basalis of Meynert & l. amygdala	Especially in Broca's and rest of frontal lobe		Limited mainly to layer II	
OTHER		Depigmented Substantia Nigra	Depigmented Substantia Nigra; Microvacuolation in frontal, temporal, & parietal lobes		

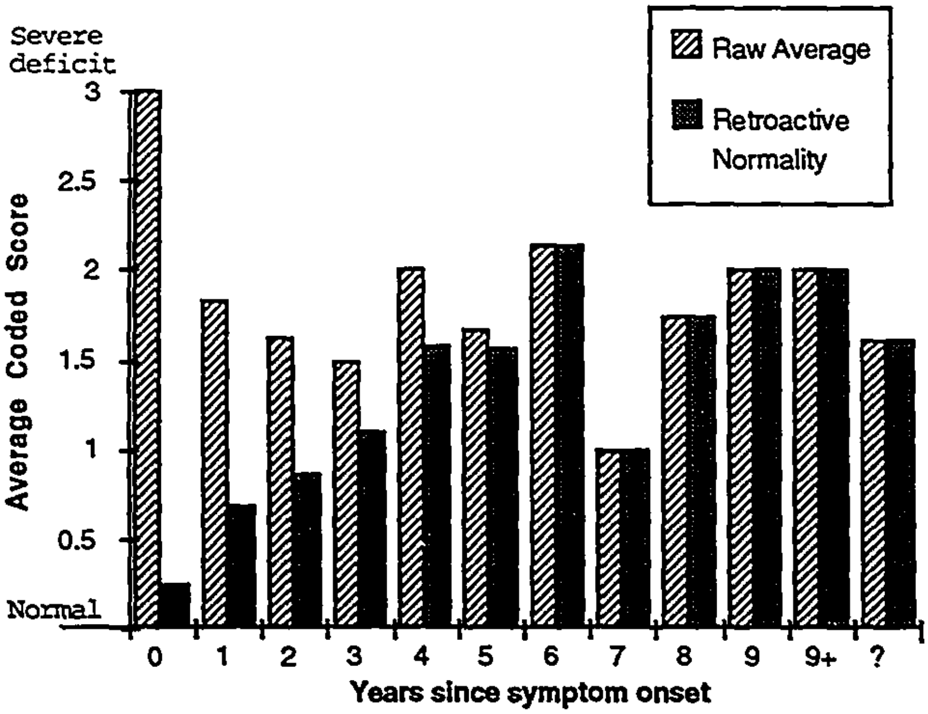
Table 3.10, Part A: Autopsy summaries of the ten patients who did not receive a diagnosis other than PPA.

SOURCE					
DURATION	5	2	9	8	5
ATROPHY	Mild: midline cerebellar vermis.	Mild /moderate focal atrophy of the l. sup. frontal gyrus	Mainly l., esp. frontal, temporal, fronto parietal & lateral parietal/occipital. Grossly atrophic hippocampus, amygdala, caudate, & putamen on l. only.	Gross bilateral atrophy, especially l. middle & sup. frontal gyri. Enlarged lateral, 3rd vent. & (mildly) 4th	Slight enlarged lateral vents. Amygdaloid nucleus atrophy.
NEURONAL LOSS			Loss of large Layer III & V pyramidal cells in frontal, fronto-parietal, & anterior temporal lobes	Loss of large pyramidal cells in frontal, fronto-parietal, & anterior temporal lobes	Loss of pyramidal cells in layer III, but not layer V
PLAQUES			No	No	
TANGLES			No	No	
SPONGIOSIS			Widespread, esp. in Layers II & III. Moderate loss in l. amygdala	Widespread, esp. in layers II & III.	Spongiosis in layer II of frontal. & temporal lobes
PICK BODIES					No
ASTRO-CYTOSIS / GLIOSIS	Limited to Layer II		Mild, even in severely affected areas	Mild, even in severely affected areas: throughout the cortical layers	Astrocytosis in layers I to III
OTHER	Microvacuolation limited to l. inf. frontal gyrus.				

Table 3.10, Part B: Autopsy summaries of the ten patients who did not receive a diagnosis other than PPA.

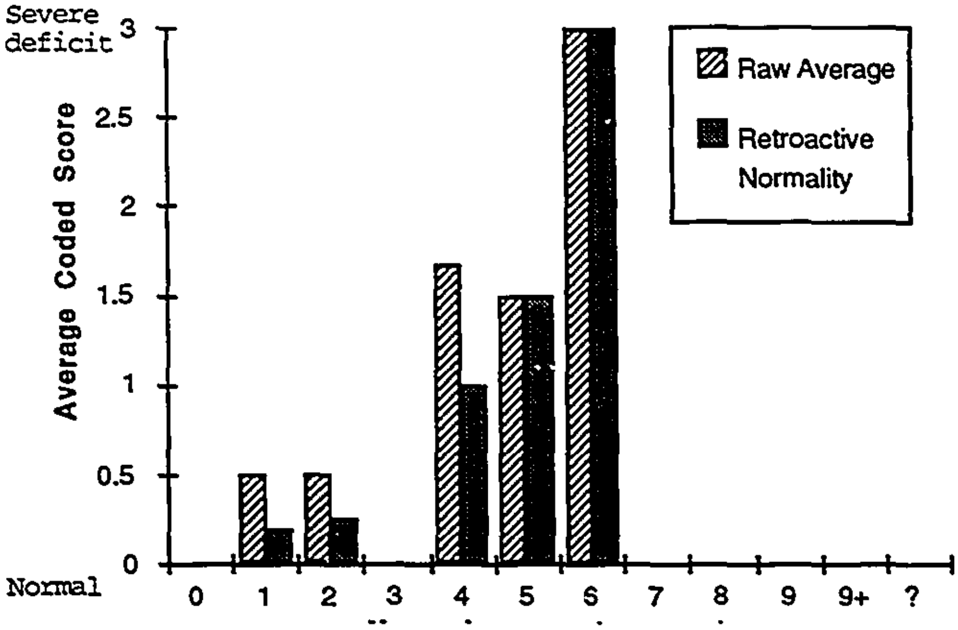
Neuropsychological Testing

Encoding all neuropsychological tests results yielded a sparse matrix. Of the 6720 cells (112 patients by 12 reporting periods by 5 test categories) only 356 (5.2%) had values. The application of the 'retroactive normality' heuristic added inferred zero values to a further 250 cells, leaving 7.6% filled. Because the data matrix is so sparse, it is not reproduced here. Instead, the average raw and normality-adjusted results for each of the five test categories are summarized graphically in Figures 3.1 through 3.5. The number of observations per reporting period and in total is reported below each of the five graphs.



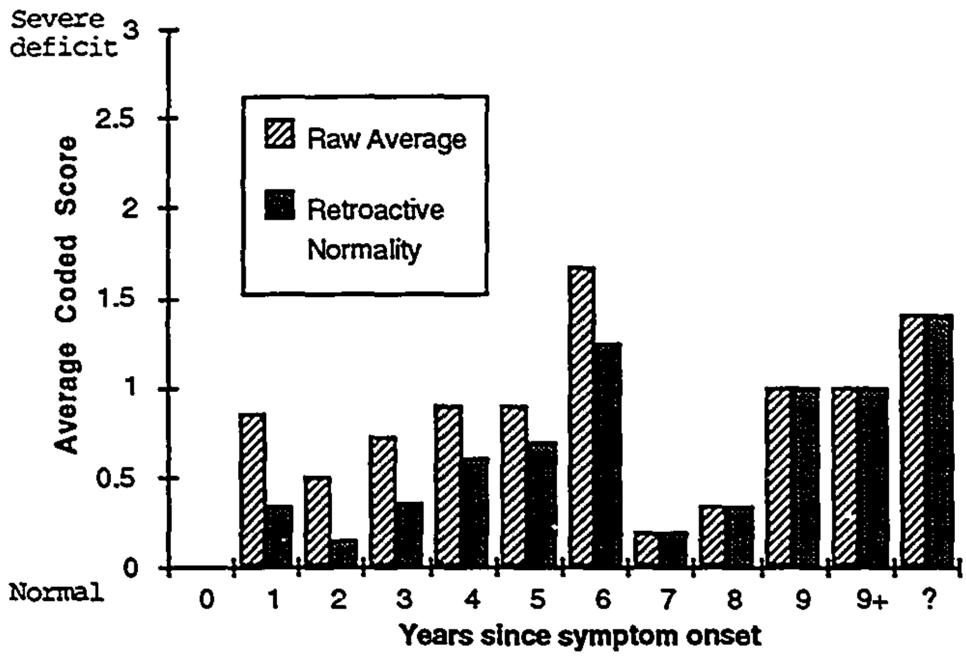
YEARS	0	1	2	3	4	5	6	7	8	9	9+	?	TOTAL
RAW N	1	6	8	14	19	15	7	5	4	2	5	26	112
RETRO N	12	16	15	19	24	16	7	5	4	2	5	26	151

Figure 3.1
Tests Of Oral Naming: Coded Results By Reporting Period



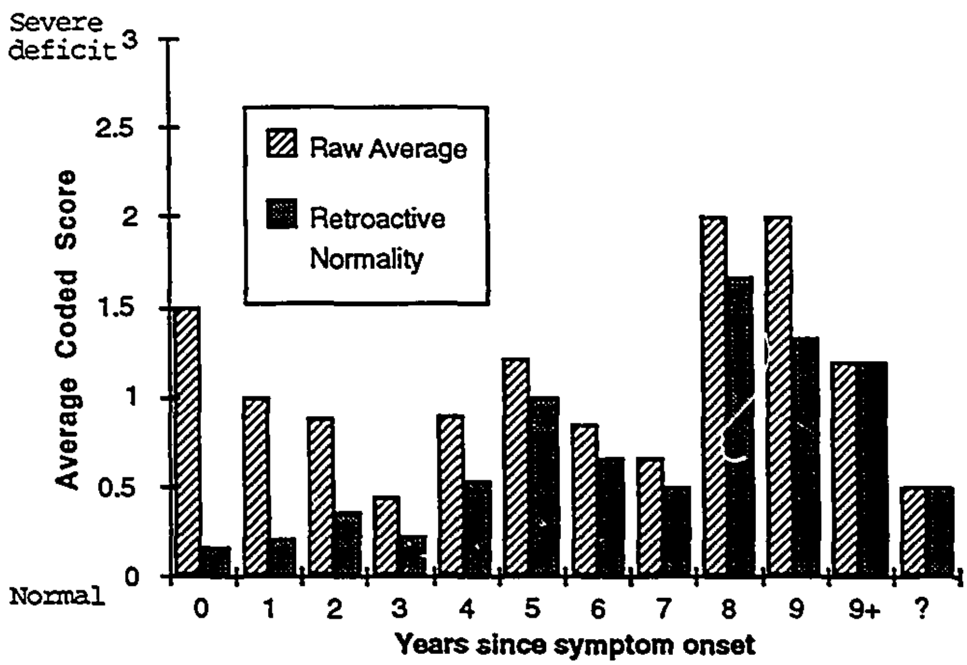
YEARS	0	1	2	3	4	5	6	7	8	9	9+	?	TOTAL
RAW N	0	2	2	0	3	4	1	0	0	0	0	0	12
RETRO N	4	5	4	2	5	4	1	0	0	0	0	0	25

Figure 3.2
Tests Of Written Naming: Coded Results By Reporting Period



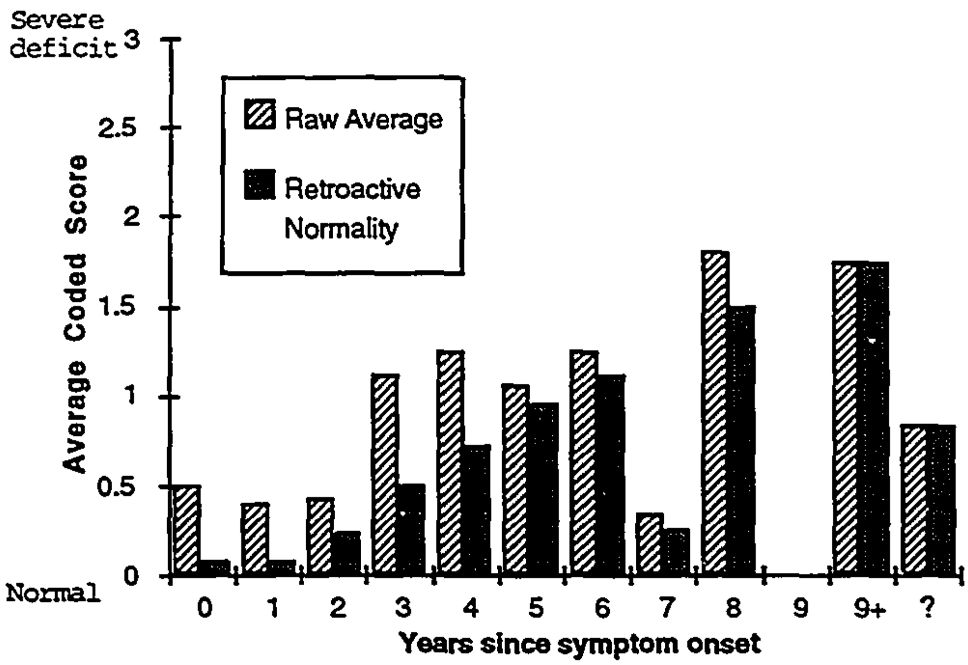
YEARS	0	1	2	3	4	5	6	7	8	9	9+	?	TOTAL
RAW N	0	7	4	7	10	10	6	5	3	2	1	5	60
RETRO N	13	18	13	14	15	13	8	5	3	2	1	5	110

Figure 3.3
Tests Of Reading: Coded Results By Reporting Period



YEARS	0	1	2	3	4	5	6	7	8	9	9+	?	TOTAL
RAW N	2	4	8	9	9	14	7	6	5	2	5	12	83
RETRO N	19	19	20	18	15	17	9	8	6	3	5	12	151

Figure 3.4
Tests Of Repetition: Coded Results By Reporting Period



YEARS	0	1	2	3	4	5	6	7	8	9	9+	?	TOTAL
RAW N	4	5	14	9	11	18	8	3	5	1	4	6	88
RETRO N	27	25	25	20	20	20	9	4	6	2	4	6	168

Figure 3.5
Tests Of Comprehension: Coded Results By Reporting Period

DISCUSSION

Demographic Data

The reported data reveal a number of interesting findings. The demographic data are surprisingly skewed towards males (by a ratio of nearly 2:1), despite the fact that females outnumber males among the elderly group from which the patients were drawn. This may reflect either a sampling bias, or a true sex difference. Although the average age at symptom onset, 59.1 years, is surprisingly low, the standard deviation (9.6 years) is quite high, reflecting the fact that PPA symptoms have been reported (Mesulam, 1982) in patients as young as 17 years.

Presenting Problem

The most commonly reported presenting symptom, a word-finding deficit, is also the most commonly documented aphasic symptom associated with DAT (Sim and Sussman, 1962; Irigaray, 1967; Barker & Lawson, 1968; Appell et al, 1982; Bayles & Tomoeda, 1983; Rosen, 1983; Cummings et al, 1985; Hier et al, 1985; Flicker et al, 1987; Murdoch et al, 1987). The fact that word-finding deficits are experienced subjectively and often reported by patients as memory deficits may account in part for the fact that PPA went undetected until it was first defined by Mesulam in 1982. Diagnosticians must carefully distinguish between naming errors and memory deficits.

Neuroanatomical data

The fact that 43% of the MRI scans reported finding evidence of anomalies in both hemispheres suggests that many cases of PPA are probably not pure, but could reasonably be expected to have deficits that are not confined exclusively to language skills.

Almost half (18 of 38, or 47%) of the scans which purported to find bilateral changes did not specify where these anomalies were. The lack of detail suggests that the bilateral scans may have shown diffuse global degeneration, lending support to the idea (discussed below in the context of autopsy data) that many patients who have been diagnosed with PPA may have been patients whose brains were generally deteriorating due to some known dementing process.

Although the PET and SPECT scans showed a lower rate of bilateral involvement, they still suggest bilateral involvement in 31% on the cases diagnosed as PPA. The frontal lobe was strongly implicated, mentioned alone or in conjunction with other sites, in 72% of the bilateral scans, and 63% of the left-sided scans. This accords with our own clinical experience that frontal lobe signs are often seen as part of PPA.

The temporal lobe was mentioned as a site of anomaly in 90% of the 40 scans reporting left-sided anomalies only, but in only 22% of the 18 scans reporting bilateral involvement, a disparity whose magnitude raises the possibility that there are two distinct processes responsible for the degenerative processes which are predominantly left and those which are bilateral.

The autopsy results summarized in Table 10 suggest that PPA is often misdiagnosed. 6 (38%) of the 16 patients who came to autopsy were shown on the basis of that autopsy to have suffered from another dementing disorder.

Neuropsychological Testing

Although comprehension deficits are also reported quite commonly, this finding is somewhat misleading and difficult to interpret, since for the purposes of this paper all tests of comprehension in any modality and at any level of receptive complexity have been lumped together. The other tests give a clearer picture. It is quite clear that the initial symptoms of PPA are extremely likely to involve verbal production. Seventy-one percent of the specified presenting symptoms involved verbal production in one way or another. In contrast, only 5% of the specified presenting symptoms involved written production, while only 3% specifically involved reading.

Perhaps it is because written production difficulties are so rarely encountered as a presenting problem that written production skills have been so rarely tested in PPA patients. Only 12 of the cases reviewed reported results from tests of written naming, just over a tenth as many

(112) as reported results from tests of oral naming. The lack of such results, and the total absence in the literature of written naming test results gathered more than five years after symptom onset, makes it impossible to be sure that written production is not differentially affected in PPA.

A fairly clear pattern emerges from the oral naming data (Figure 3.1). Oral naming is increasingly affected for up to five years year after symptom onset, after which only a single result (of 39 results coded during or after the fifth year post-onset) showed no deficit at all. This suggests that the oral naming deficit in PPA peaks will certainly be manifest after five years. At that time the average deficit is not profound, with an average coded score for an oral naming deficit in each year after the fifth year of less than two.

Figure 3.6 breaks down the results by years from onset and proportional representation of each code in order to make the pattern of the deficit clearer, although representing the data proportionally masks the progressive nature of the disorder, and renders data from the first few years unreliable, since the number of measures obtained from those years is small. Despite these limitations, the graph clearly shows that by the third year the distribution of deficits is relatively constant. From that time on, a large proportion (roughly 45%) of the test results indicate the presence of a severe naming deficit. About half that many results indicate no deficit at all- although the application of the 'retroactive normality' rule would raise that proportion. A slightly larger number of tests results (about 30%) show a mild deficit. Relatively few patients score a deficit in the 'moderately impaired' range. Moreover, the proportion of patients who do does not increase with time. This may indicate that the degeneration is quite rapid once it begins, so that a random probe (i.e. the administration of a naming test) will be unlikely to catch the process in progress.

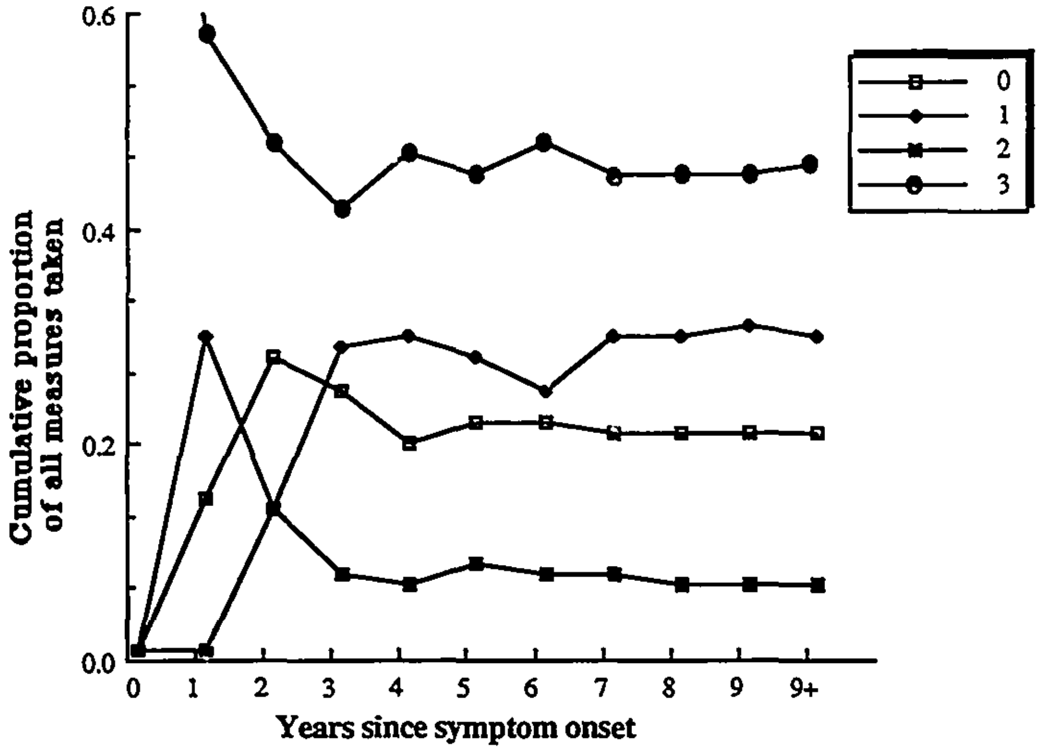


Figure 3.6:
Oral naming: Cumulative proportion of scores by code
and years since symptom onsets
Codes run from 0 (Normal) to 3 (Severely Impaired)
(Data from the first few years is unreliable,
since the number of measures on those years is small.)

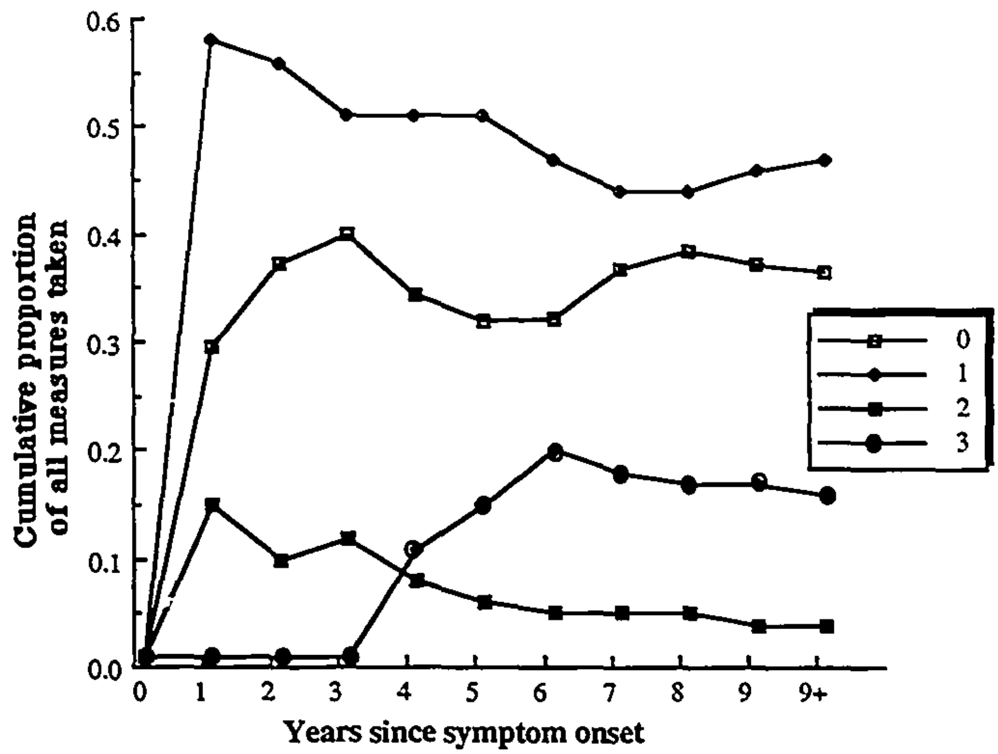


Figure 3.7:
Reading: Cumulative proportion of scores by code
and years since symptom onset
Codes run from 0 (Normal) to 3 (Severely Impaired)
(Data from the first few years is unreliable,
since the number of measures in those years is small.)

Figure 3.3 shows a relatively low average reading deficit throughout the reporting period. The breakdown of the results of the reading test (Figure 3.7) confirms that reading is relatively intact in many PPA patients. Although a majority (ranging from just over half in the second and third years post-onset, to just under half by the seventh and eighth years) of tests uncover a mild deficit, a relatively large and constant proportion (roughly 35 - 40%) of test results show no deficit at all. The proportion of results indicating a severe deficit rises quite rapidly from the third to the sixth year post-onset (at approximately the same rate as the proportion indicating a moderate deficit decreases), but never rises above a peak of about 20% in the sixth year. Severe reading deficits in PPA rarely appear before the first four or five years after symptom onset.

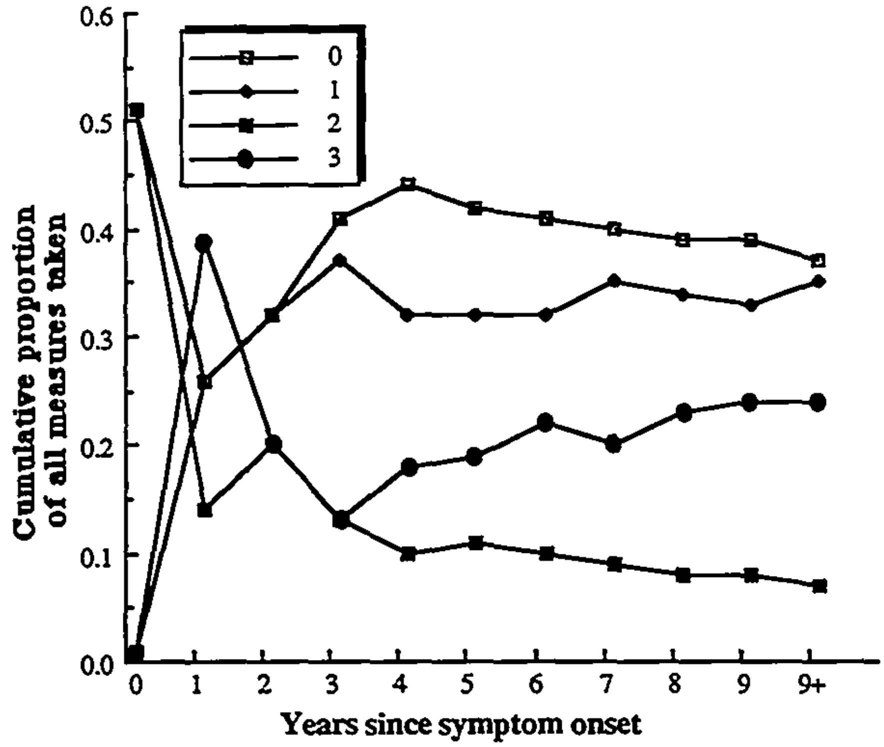


Figure 3.8:
Repetition: Cumulative proportion of scores by code and years since symptom onset
Codes run from 0 (Normal) to 3 (Severely Impaired)
(Data from the first few years is unreliable, since the number of measures in those years is small.)

The repetition results (Figures 3.4 and 3.8) are somewhat similar to the reading results. Again, the group averages (Figure 3.4) suggest that there is a minimal deficit in repetition in the early years, although by the eighth and ninth year, the average effect seems to be that of a moderate deficit. As the break-down in Figure 3.8 shows, the large majority of tests show either no deficit (roughly 40%) or a mild deficit only (roughly 35%) throughout the entire reporting period. The rising average deficit in Figure 3.4 reflects the steady rise of severe repetition deficits beginning in the third year and continuing throughout the reporting period. Such severe deficits account for just over 10% of all results in the third year, and about 20% by the ninth year. In the same time span, there is concomitant decrease in the number of moderate deficits and in the number of tests which uncover no repetition deficit at all.

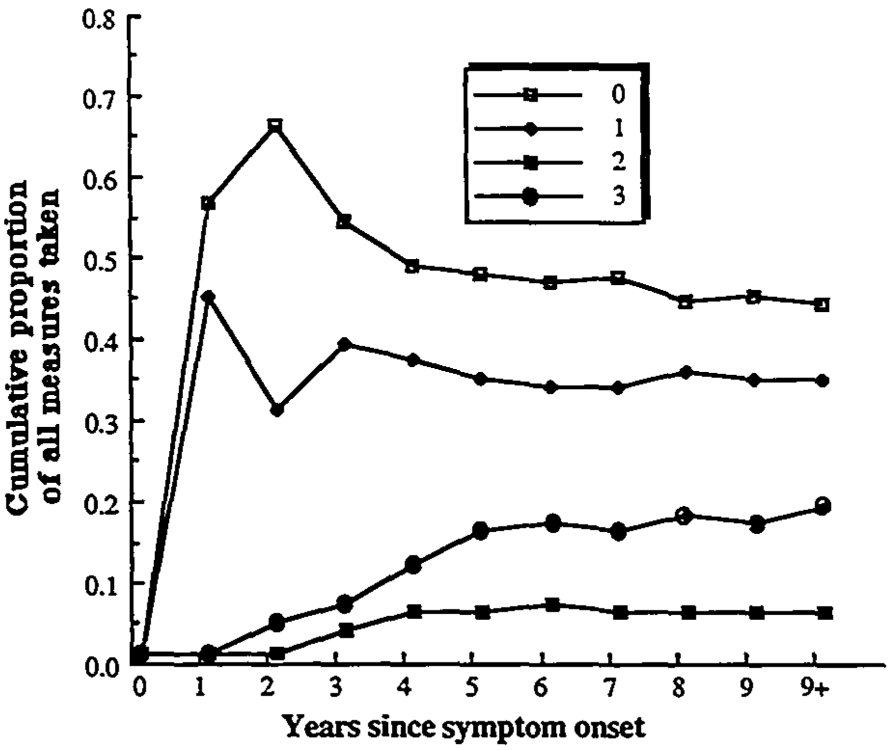


Figure 3.9:
Comprehension: Cumulative proportion of scores by code
and years since symptom onsets
Codes run from 0 (Normal) to 3 (Severely Impaired)
(Data from the first few years is unreliable,
since the number of measures in those years is small.)

Although the comprehension data are particularly difficult to interpret because the term is used so loosely in the literature, the averages graphed in Figure 3.5 do suggest that there is a general increase in comprehension difficulties which rises quite rapidly from the first year post-onset. The breakdown by severity and time in Figure 3.9 reveals a slow but steady decline in the proportion of measures which uncover no deficit or a mild deficit only, and a slow but steady increase in the proportion of measures which uncover a severe deficit. By the ninth year almost 20% of all measures reveal a severe comprehension deficit (up from less than 5% at year two), while about 50% find no deficit (down from about 70% at year two). Mild deficits are found about 40% of the time throughout the measuring period, although the proportion declines very slowly during that time.

There are many difficulties in the interpretation of this data. Severe symptoms often degenerate into untestable symptoms, so that severe pathology will be under-represented in later years. Mild pathology or lack of pathology may be under-reported, for different reasons, in early years. Moreover, our results are by test result, not patient, so there is a mixture of within-patient and between-patient data in these graphs.

CONCLUSION

Although this review makes clear the great heterogeneity in the symptoms of patients diagnosed with PPA, some tentative general conclusions may drawn.

The variability in the pattern of neuropsychological deficits reflects the variability in the underlying pathology of PPA. However, it is clear that naming is the most common and earliest deficit. By the third year about 45% of all tests of PPA patients have a severe naming deficit, while about 30% have a mild deficit. The majority of the remainder show no deficit at all. Reading deficits, if they appear at all, are likely to appear later, with severe deficits rarely seen before about the fourth or fifth years after symptom onset, and never seen in about 80% of cases. Most PPA patients never show more than a mild reading deficit. Repetition is affected at about the same rate, and in approximately the same manner, as reading, with severe deficits likely to be seen in only about 20% of cases by the ninth year. Many patients show no comprehension deficits, especially in the first two years, but the numbers showing a severe deficit rise from the first year after onset, to a peak of about 20% by the ninth year.

Forty-three percent of MRI scans and 31% of PET scans found evidence of bilateral cerebral anomalies. Differences in the distribution of anomalies between the groups with bilateral and unilateral changes suggest that there may be two separate processes involved.

Chapter Four: PPA Analysis

Any fool can make a tidy map;
the trick is to make an accurate one.

Patricia Churchland & Paul Sejnowski
The Computational Brain

In serial case studies such as this one, it is customary to present the raw results for the individual cases in the body of the study. It is important that these details be included to assist the reader, for three closely related reasons:

i.) Because a defining idea behind the case study approach in cognitive neuropsychology is that it is necessary to treat each patient as an individual study (see Chapter 1);

ii.) Because many of the conclusions that can be drawn depend upon an unambiguous and fine-grained definition of the individual deficits;

iii.) Perhaps most importantly, because the definition of complex neuropsychological deficits necessarily involves a theoretical analysis, rather than simple presentation, of the data.

However, this thesis is unusual both in the scope of the individual cases presented, and in the atypically large number of individual cases to be analyzed. These two features make it impractical to include the individual cases in the body of the study. Those individual studies are therefore presented in Appendix B, to which the reader may refer for the definitions of the high level deficits which are summarized and analyzed in this chapter. Appendix B also includes neuroanatomical data from PET, MRI, and CAT scans for many patients.

The language deficits which are detailed in Appendix B are summarized in Table 4.1, which highlights the heterogeneity among the deficits of patients diagnosed with PPA. In this chapter I will attempt to systematically analyze this heterogeneous set of symptoms in an attempt to formulate a coherent understanding of the symptoms. The data set will be approached in two ways. In the first section of this chapter, I use measures of dissimilarity between the patients to try to re-organize the data set in Table 4.1. This is a risky approach, as it tempts us towards the 'statistical co-occurrence' methodology which was dismissed as an historical dead end in Chapter 1. However, so long as we do not take their descriptive utility to be explanatory, analyses based upon simple co-occurrence can be a useful tool for uncovering correlations of symptoms which may later be found to have a solid theoretical basis. In the second section of this chapter, I approach the data using methods which are more in keeping with the methodology of modern cognitive psychology, focusing upon four main symptom sets which may be reasonably believed to be theoretically coherent according to modern models of language processing.

Subject	AB	DM	JD	CD	JH	BH	BL	JL	CM	ES	MW
Auditory Input deficit	Y		Y	(Y)	Y	(Y)	(Y)				
Reading Deficit		Y	Y	(Y)	Y	Y	Y			Y	(Y)
Anomia				(N)	Y			-		Y	
Agnosia	Y	Y					Y	Y	Y		(Y)
Objects < Organic							Y		Y		
Organic < Objects	Y										
Bad Abstract Words				(Y)	Y			Y	Y	Y	Y
Abstract word sparing		Y									(O)
Morphology Deficits											
Production	Y	(Y)	Y	Y	Y	(W)	Y	Y	(Y)	Y	(Y)
Comprehension	Y		Y	Y	Y			Y		Y	Y
Sentence Deficits											
Production		Y		(O)	Y			Y	R	R	Y
Comprehension	W		Y	(N)	(Y)			Y	Y	(Y)	(Y)

Table 4.1: Summary Of PPA Deficit Identified With PAL
Y = Deficit is present; N (or blank) = Deficit is absent; - = Deficit is unidentifiable
Deficits in parenthesis are equivocal.
R = Reversible sentences only (Syntax Deficit)
O = Oral Modality Only; W = Written Modality Only

GLOBAL OVERVIEW

In order to clarify the pattern underlying the deficits a number of heuristics were applied to the data. First, the data presented in Table 4.1 was recoded as binary data, using a '1' to signal the presence of a deficit, and a '0' to signify its absence. As a simplifying step, deficits which were equivocally present were recoded as '1' and deficits which were equivocally absent were recoded as '0'. In those cases where a deficit had been identified in a single modality (abstract word sparing, sentence comprehension, and sentence production), the results were divided by modality, creating two categories from each of these three individual categories. Patient performance in each of the six new categories was encoded with a single binary digit. The binary encodings allowed for calculation of the Euclidean distance between all possible pairs of patients, as a measure of their similarity. Similarity measures were used as input to a multi-dimensional scaling package (Systat version 5.2.1, using the Kruskal loss function with linear regression) which generated a scaling in two dimensions. The scaling stabilized after only 24 generations, the Shephard diagram was smooth, and the resultant diagram accounted for 98.4% of the variance. From this scaling, four main clusters (themselves clusterable into two main groups) were identified visually, with one individual left out. The scaling, along with the identified clusters, is reproduced in Figure 4.12.

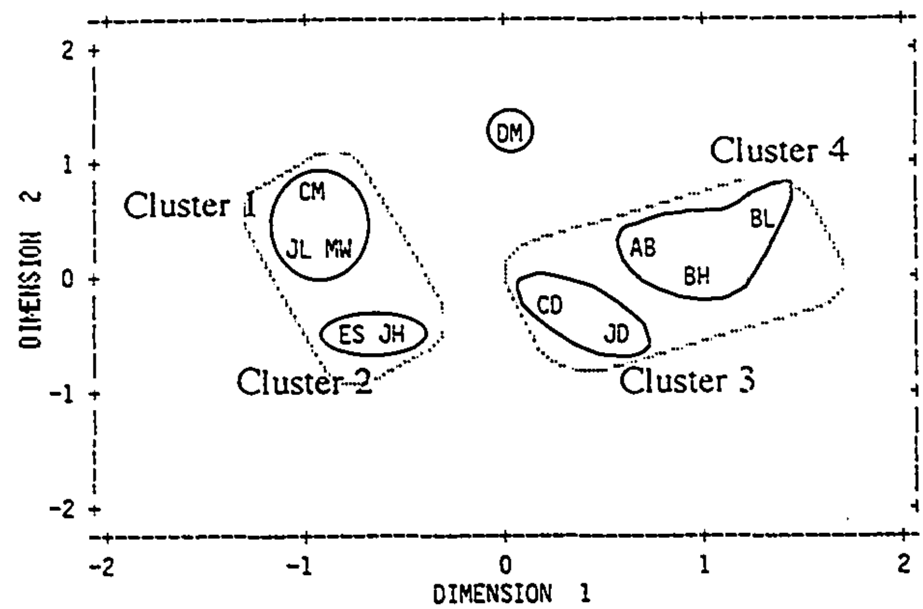


Figure 4.12
Multi-Dimensional Scaling
By Euclidean Distance In 'Symptom Space'
(see Table 4.1 and accompanying text)

The number of simplifying measures, approximations, and subjective decisions necessitated by this approach make it difficult to assess its utility objectively. However, the utility of such an heuristically-guided approach to pattern recognition can be gauged by its fruitfulness in reorganizing the original data. The data presented above in Table 4.1 is re-presented (in its binary encoded form) in Table 4.2, where it is re-organized in terms of the clusters identified on the multi-dimensional clustering. In this form the regularities which underlie the clusters may be readily identified.

	Group A					Group B						
	Cluster 1			Cluster 2		Cluster 3		Cluster 4			Cluster 5	
	MW	CM	JL	ES	JH	JD	CD	BH	BL	AB	DM	Total
Auditory Input deficit	0	0	0	0	1	1	1	(1)	(1)	1	0	6
Reading Deficit	(1)	0	0	1	1	1	(1)	1	1	0	1	8
Anomia	0	0	0	1	1	0	0	0	0	0	0	5
Agnosia	1	1	1	0	0	0	0	0	1	1	1	6
Organic < Object	0	1	0	0	0	0	0	0	0	1	1	3
Object < Organic	0	0	0	0	0	0	0	0	1	0	0	1
Bad Abstract Words	1	1	1	1	1	0	(1)	0	0	0	0	6
Abstract word sparing	1	0	0	0	0	0	0	0	0	0	1	2
Bad Morphological Production	(1)	(1)	1	1	1	1	1	(1)	1	1	(1)	11
Bad Morphological Comprehension	1	0	1	1	1	1	1	0	0	1	0	7
Bad Sentence Comprehension (O)	1	1	1	1	1	0	1	0	0	0	1	7
Bad Sentence Comprehension (W)	1	1	1	1	1	0	0	0	0	0	1	6
Bad Sentence Production (O)	1	(1)	1	(1)	1	1	(0)	0	0	0	0	6
Bad Sentence Production (W)	(1)	1	1	(1)	(1)	1	(0)	0	0	(1)	0	7
Total Deficits	10	8	8	9	10	6	6	3	5	6	7	

Figure 4.2
PPA Patients By Groups

Clusters 1 & 2

We may consider clusters 1 and 2 together, since the two clusters are close together and clearly separate from the rest of the patients. Cluster 1 contains three patients: ES, MW, and JL, while cluster 2 contains ES and JH. The main characteristic which links all of these subjects together is the severity of their symptoms. On average, the five subjects in the two cluster had 9.0 of the 14 characters (symptom sets) which were identified, compared to an average of 5.7 among the remaining 6 subjects.

The pattern of symptom distribution in these two clusters is quite clear. All five patients were either anomic or agnosic. All five patients scored significantly low on abstract word comprehension, production of affixed words, sentence comprehension, and sentence production. All but one (CM) also scored significantly low on affixed word comprehension. CM was also the only patient of the five to show any category-specific deficits.

The difference between the two clusters is also clear. The three patients in cluster 1 were agnosic and showed no (or, in MW's case, only weak) signs of reading difficulties, while the two in cluster 2 (ES and JH) were both anomic, and had reading difficulties. They were also the only two patients who presented with an almost total disruption of productive output.

It is of interest to note that only one (JH) of the five patients in these clusters showed signs of the auditory input deficit which was identified in five of the six remaining patients. Because it is seen among the most severely afflicted patients, this apparently systematic sparing suggests the possibility that the five patients may in fact constitute a true subtype of PPA, rather than being united only by the severity of their symptoms. However, note that there is no apparent theoretically-motivated connection between either the spared or impaired symptoms.

The fact that the five most severely affected patients are also the only five patients with an unequivocal abstract word deficit is also intriguing. It is difficult to understand this without more information about the relation between abstract word comprehension and performance on other language tests. It may that the comprehension deficit appears late in the disease process, and thus is likely to be seen mainly in severe cases which have many other deficits as well. Alternatively, it may be that there is a causal relation between an abstract word deficit and failure on the morphology and sentence processing tests, so that the fact of having such a comprehension deficit is sufficient to guarantee failure on a myriad of other tests.

Clusters 3 & 4

Like clusters 1 and 2, clusters 3 & 4 contain five patients who are very similar to each other. All five patients (and only one patient outside of this group) showed signs of an auditory input deficit and a reading deficit. None of the five (and all but one outside of this group) showed sentence comprehension deficits.

JD and CD are grouped together in the third cluster. They differed from the patients in the first cluster mainly in having no naming difficulties. Like the five patients in the first cluster, they were clearly impaired at morphological production and comprehension.

The fourth cluster may be a 'precursor' to the third, as the three patients in the fourth cluster (BH, BL, and AB) differed from the two in the second mainly insofar as they were less impaired. Morphological comprehension was preserved in two (BH and BL) of the three patients, but morphological production was impaired in all three. AB also was the only patient of the three with signs of agnosia or of difficulty in processing sentences.

It is particularly appropriate for BH and BL to be grouped together, since (as chance, rather than the non-specific method of quantifying the deficits, would have it) their impairments in auditory processing were also qualitatively similar, and quite different from the auditory input deficits identified in the four other patients who manifested such a deficit. Although BL and BH were both impaired on the phoneme discrimination task, neither one also showed the repetition impairment which is most clearly suggestive of a deficit in auditory processing. Their deficits were inferred in both cases from the existence of a modality-specific comprehension deficit. It is possible that the comprehension tasks are sensitive to early impairment of the auditory processing system.

Only three patients showed category-specific effects in naming. BL was the only one of these three to show a category-specific impairment for inorganic objects.

Patient DM

DM is unique not only because of his abstract word sparing (seen, equivocally, in only one other patient, and unusual enough to be worthy of its own category) but also because he is the only patient who was impaired at sentence comprehension without being impaired at sentence production. He was also one of only two patients who showed signs of a category-specific deficit limited to organic stimuli.

It should be noted that the scaling in Figure 4.12 captures the referral histories of the patients almost perfectly. The five patients in Clusters 1 and 2 were all referred by a single neurologist. The two patients in Cluster 3 were referred by another neurologist. Two of the three patients in Cluster 4 (BH and BL) were referred by a third neurologist. AB and DM were individual referrals from a fourth and fifth neurologist. In light of this close mapping between similarity and referring source, we cannot exclude the worrisome possibility that the clusters are capturing something about each neurologist's concept of PPA, rather than something about PPA per se. This possibility emphasizes the need for systematic studies such as these, which may eventually serve to unify the diagnostic category of PPA in a systematic manner. I will return to this topic in my final chapter.

ANALYSIS BY DEFICIT

The groupings above are useful for generating questions about the nature of PPA and for providing us with a rough global understanding of how language deficits manifest themselves in the eleven patients of this study. However, they mask a great deal of detail, ignoring both differences between patients within a single group, and similarities between patients in different groups. Moreover, as mentioned above, they may also tempt us to give undue weight to random correlations of deficits in the data which have no theoretical importance. In this section, I focus attention on a closer analysis of four important classes of deficits which span across all the clusters: the auditory input deficits, the naming deficits, the reading deficits, and the hitherto-undocumented deficits in affixed word production and comprehension. As well as trying make the description of the impairments in PPA more precise, my purpose in doing so is to outline precisely why taxonomic exercises in aphasics are so difficult, by underscoring the flexible nature of the groupings, and by detailing the important variation which can exist even within the 'same' impairment. I will also spell out the meaning of some important patterns in the impairments which have general theoretical implications.

i.) Auditory Comprehension Deficits

Six of the eleven patients in this study (AB, JD, JH, CD, BH and BL) had a set of deficits which indicated that they were having particular trouble recognizing auditory input. Note that five of these patients (all except JH) were similar enough to each other in other ways to cluster together into one of the two 'super-clusters' identified above. Patients with such a deficit are sometimes referred to as having 'auditory agnosia'. The term is slightly inaccurate, because it blurs the distinction drawn in Chapter 2 between semantically-mediated word access, impairment of which constitutes a true auditory agnosia, and nonsemantically-mediated word access, impairment of which constitutes an auditory-specific comprehension deficit which need not be modulated by semantic access. In extreme cases of auditory comprehension deficits, the more accurate term 'pure word deafness' appears preferable. For the sake of clarity, I will try to avoid the fine semantic distinctions in this section by referring only to problems of auditory input comprehension.

In order to be able to recognize the problem as a specific failure of auditory input comprehension, one of two kinds of evidence was needed, either:

i.) Either failure of phoneme discrimination and repetition, with comparatively spared performance on tests of oral production (such as reading) which were not mediated through auditory input, or ii.) Consistent dissociations between analogous tests of auditory and written comprehension.

The relevant results are presented in Table 4.3. As that table makes clear, no patient met the most stringent possible criteria, namely, that there be no evidence of a reading deficit at all, and a consistent difference in favour of written input on all tests which were comparable across modalities. In part, this reflects the sensitivity of the reading test (on which almost every patient showed deficits) and/or the ubiquity of reading deficits in PPA.

AB has the purest case of auditory agnosia. She was the only patient who showed a difference in favour of reading between the normalized scores for reading and repetition. She was almost perfectly consistent in her performance pattern in favour of the written modality, scoring slightly higher in the auditory modality on only one test (lexical decision). She also showed significant frequency and lexicality effects (the latter debatable due to the floor effect) on the repetition test but not the reading test, and had the best reading score (91%, about five SDs below norms) of the five patients considered in this section. Few diagnosticians would dispute that AB's symptoms are compatible with a diagnosis of 'pure word deafness'.

Every other patient with an identified auditory word comprehension deficit has accompanying deficits which obscure the overall picture.

JH is perhaps the second clearest case, although he had deficits in both reading and abstract word comprehension complicate our understanding of his auditory comprehension deficit. JH had the largest and most consistent differences between auditory and non-auditory tests, being far better (at least 9 standard scores) in the written modality on lexical decision, word-picture matching, and lexical decision of derived words, and slightly better (0.5 standard scores) on affixed word picture matching. His reading was anomalous among the six patients with auditory input deficits inasmuch as he was significantly better at reading nonwords than words.

JD had an auditory comprehension deficit which is similar to, but less pronounced than, JH's deficit. He also had great difficulty reading. He was the only patient who showed a consistent advantage of written over auditory stimuli across all five tests for which the comparison is possible. However, his advantages were comparatively small on both the written word picture matching tests, and on the test of lexical decision of affixed words, reflecting an apparent deficit with affixed word comprehension. The difference between the two modalities on the abstract word comprehension tests and the lexical decision tests was more pronounced.

CD is a milder case, although her performance was still quite impaired. She showed small differences in modalities on four of the five cross-modality tests (all except word-picture matching of affixed words, once again reflecting a difficulty with affixed words). She was massively impaired at repeating nonwords (20%) and performed at chance in identifying differences between two phonemes.

BH and BL have auditory comprehension deficits which were similar to each other's deficits, but very different from the other deficits. Neither one scored significantly low at repetition, but both showed auditory-specific deficits on all cross-modality tests except lexical decision.

It is interesting to note that the largest difference between modalities was seen between the two abstract word tests in every patient for whom the comparison was possible. In both of the remaining patients (AB and JH) there was also some evidence of a modality-specific abstract word comprehension deficit, the extent of which is unquantifiable only because they were both unable to complete the auditory version of the test. By comparison, only one (DM) of the five PPA patients who did not show an auditory comprehension deficit scored lower on auditory comprehension of abstract words than on written comprehension of abstract words. This consistency suggests the possibility that the abstract word comprehension test is particularly sensitive to auditory comprehension deficits.

It is difficult to assess this hypothesis further with the present data set. Phoneme discrimination impairments, which constitute a necessary but not sufficient symptom for identifying acoustic comprehension deficits, can arise from lesions "over a wide area of the left hemisphere" (McCarthy and Warrington, 1990a, pp. 139) and "There is only limited evidence from group studies on the localizing significance of word-comprehension deficits" of any kind (McCarthy and Warrington, 1990a, pp. 140). Almost nothing is known about the localizing significance of abstract word comprehension deficits in particular, or about the general relation between disturbances in acoustic and phonetic processing and comprehension deficits (Caplan, 1992).

Subjects	AB	CD	JD	JH	BH	BL
Low PD	Y	Y	Y	Y	Y	Y
Same > Diff.	-	Y	Y		Y	
Low Repetition	Y	Y	Y	Y		
W > NW	Y	Y	Y	Y		
HF > LF	Y	Y		Y		
Low Reading	Y	Y	Y	Y	Y	Y
W > NW			Y	(Y)	Y	Y
HF > LF		Y		Y		
Rep - Reading	**4.6**	-2.3	-25.1	-15.9	-14.8	-7.1
Auditory - Written Standard Scores Below Normal						
Abstract Word Comp.	**Y**	**1.6**	**5.2**	**?**	**4.1**	**9.2**
LD	-0.3	**1.0**	**3.4**	**13.4**	-1.9	-1.9
WPM	**4.9**	**0.8**	**0.1**	**9.8**	**1.2**	**5.6**
LD (Affixed)	**1.4**	**0.4**	**0.3**	**16.6**	**0.3**	**1.9**
WPM (Affixed)	**0.7**	-4.5	**0.7**	0.5	3.1	**3.9**

Table 4.3:
Results For Patients With Evidence Of Auditory Processing Deficits
A 'Y' indicates the presence of an effect at the P <= 0.05 level.
Brackets indicate that the effect was in the opposite direction to the label.
Difference scores are in standardized units (z-scores).
Bold type indicates deficits in the direction expected (most often, a written advantage).

ii.) Naming Deficits

There are two non-mutually-exclusive main types of naming deficits, reflecting the fact that naming is a mapping of information from two domains: a lexicon and a semantic storage (McCarthy and Warrington, 1990a, Chapter 2; McCarthy & Warrington, 1990b, Shallice, 1988). Failure to access the lexicon will result in naming difficulties with intact access of semantic information- an anomia. Failure to access semantic storage will result in naming deficits with associated failure on purely semantic tasks- an agnosia. In this thesis, I have taken a naming deficit which is distinguishable along semantic lines (a naming deficit with evidence of category specificity) to be indicative of an agnosia, since semantic characteristics belong to the object rather than to the object's name. It must be admitted, however, that this may be something of a simplification, since it is possible that object labels are grouped semantically in the brain, or that there is some more subtle connection between object labels and semantic categories (Dixon, Bub & Arguin, 1994).

Distinguishing between anomia and agnosia is a difficult, and sometimes impossible, task, especially in a patient who has other deficits, and with the limited tests of semantic access upon which I relied. However, it is important to make an effort to distinguish them. As was first suggested in 1890 by Lissauer (McCarthy, 1975), semantic access and lexical access are now clearly known to be modulated by very different brain regions (possibly even different hemispheres- see McCarthy & Warrington, 1990a) and have been clearly shown to be functionally distinct (McCarthy and Warrington, 1990a; McCarthy & Warrington, 1990b).

Eight of the eleven PPA patients had naming deficits of some kind. Only two patients (ES and JH) may be considered unequivocally anomic. These two patients are extremely similar in other important ways: their test profiles are almost identical, resulting in their being grouped together in the multi-dimensional scaling. Both patients presented with an almost complete inability to speak. The remaining six patients (BL, AB, DM, MW, JL, and CM) showed signs of impaired semantic access that suggest that their naming difficulty is more likely attributable to an agnosia. The results which are relevant to this distinction are presented in Table 4.4.

ES is the clearest case of a pure anomia. She achieved very poor results on tests of written and oral naming, but was within the normal range on both the semantics battery and written word-picture mapping, and just one item below the cut-off point at auditory word-picture mapping. There was no evidence of a category effect in any one of the naming or picture-matching tests.

JH is also anomic. However, his case is complicated by his auditory impairment, which may well have affected his performance on the auditory word-picture matching test and by a consistent reading deficit which especially affected his reading of long words. He was in the normal range on the semantics battery, and much better at written word-picture mapping than written naming, both signs suggesting that his problem lies in accessing the lexicon rather than in accessing semantics.

	Anomia		Agnosia					
Patient	ES	JH	AB	BL	CM	DM	JL	MW
AWP	2	14	5	6		9	4	21
WWP		4			2	6	5	10
O-NAME	10	16	3	6	10	14	19	19
W-NAME	40	43	6	27	17	42	-	50
SEM.				3	3	3	5	3
CATEGORY SPECIFICITY?								
AWP			IN<ORG	IN < FV AN < FV	FV<AN	FV < IN		-
O-NAME			IN<ORG	IN < FV AN < FV		-	-	-
WWP			IN<ORG		FV<OB	FV < AN FV < IN		-
W-NAME			IN<ORG	IN < FV	FV < AN FV<OB	AN < IN FV < AN	N/A	-

Table 4.4:
Results For Patients With Evidence Of Naming Deficits
Only significantly low scores are reported.
AWP = Auditory word-picture matching
WWP = Written word-picture matching
O-NAME = Oral picture naming
W-NAME = Written picture naming
SEM. = Global semantics battery score
Numbers are standard scores below average.
Reported dissociations use a 2 SD cut-off point.
Dashes indicate that the score was too low to allow for a dissociation.
'N/A' indicates the test was not administered.

It is extremely interesting to note that both JH and, to a lesser extent, ES (two patients who, as noted above, are similar in so many other ways) showed regularities between their lexical processing deficits and their anomic symptoms. ES had particular difficulty reading (and also repeating) both short words and short nonwords, and also showed a fairly large effect of name length in the same direction on her oral naming test. The consistency was statistically stronger in JH's case: he had significantly more difficulty reading (and making lexical decisions about) long words than short words, and also showed a significant effect of name length in the same direction on the test of written word picture matching. Note that the consistencies go in both directions (that is, differentially impair words in each of the two length categories). These observations suggest a possible extension to the view (expounded by Patterson & Hodges, 1992 and Graham, Hodges & Patterson, 1994) that anomic symptoms may be related to deficits in processing lexical characteristics in reading, an hypothesis which will be considered in more detail in the next section. Graham et al

(1994) documented similar consistencies between comprehension and reading in three PPA patients, looking at word frequency only. The implication of the present data is that anomic symptoms may be more generally sensitive to lexical characteristics: in this case, to word length. Note, however, that neither ES nor JH shows a consistent effect of frequency in the tests of lexical access and naming.

BL, AB, CM, and DM, the four clearest agnosics, all showed signs of a category specific deficit.

AB was consistently better at naming or matching names of organic than inorganic stimuli, the only such patient to show a dissociation in that direction. Given the inorganic specificity of her deficit, it is not surprising to see that she did not score below norms on the semantics battery, which includes only organic stimuli. We may discount her poor performance on auditory word-picture mapping as being due to her general deficit in auditory processing.

All of the remaining patients classified as agnosic scored below norms on the semantics battery.

BL showed evidence of a category-specific sparing for fruits and vegetables in auditory word-picture matching, and oral naming. In written naming she was significantly better at naming fruits and vegetables than objects, but not animals. Her written word-picture matching was too good to show any significant category effects.

CM showed the opposite effect: a specific deficit limited to fruits and vegetables. He was significantly worse with fruit and vegetable stimuli than animal stimuli on the auditory word-picture matching test and the written naming test, and significantly worse with fruits and vegetables than objects on the written word-picture matching test and the written naming test.

DM showed a similar pattern, being significantly worse with fruits and vegetables than objects in both the word-picture matching tests, and significantly worse with fruits and vegetables than animals in the written naming tests.

It is more difficult to decide whether MW and JL have difficulty accessing words or semantics, since both of them were either unable to complete or scored significantly low on the naming tests, the word-picture matching tests, and the semantics battery. Neither of them showed dissociations along the lines of semantic categories, mainly because their scores were too low globally to allow for a meaningful analysis of the results by category.

The localizing value of these findings is again uncertain, for very similar reasons to the two reasons presented in the last section: in part because there is only cursory information about the exact nature of the semantic deficits in these patients, and in part because the nature of the lesion which produces such deficits is not well understood. Basso, Capitani, & Laiacona (1988), reporting on a patient with symptoms very much like several of the cases reported here (with a progressive language impairment, no dementia, and a category-specific agnosia limited to organic stimuli) reported that the lesion was in the inferior perisylvian region.

iii.) Reading Deficits

In light of Patterson and Hodges' recent findings (Patterson & Hodges, 1992; Graham et al, 1994) which suggest that there is a close connection between semantic deficits and reading deficits, it is appropriate to turn our attention now from a consideration of the former to a consideration of the latter, and to the link between the two.

Six patients (ES, JH, JD, BH, and BL) had unequivocal reading deficits. More equivocal mild deficits were found in three other patients (AB, MW, who each made only four reading errors, and CD, whose errors showed a dubious pattern, as described above). The relevant results are summarized in Table 4.5.

	JH	ES	DM	BL	AB	MW	JD	BH	CD
Reading	22	20	15	6	4	4	33	18	6
W > NW	(Y)	Y	Y	Y		(Y)	Y	Y	
HF > LF	Y	Y							Y
HI > LI			(Y)				Y	(Y)	(Y)
S > L	Y	Y	Y				Y	Y	Y
REG > IRR	Y	Y	Y		Y		(Y)		Y
PSIM > PCOM									
OSIM > OCOM	Y	(Y)		Y	Y			Y	
Error Counts:									
Semantic	0	0	0	0	1	0	0	0	0
Phonological	13	10	6	6	0	1	18	12	1
Regularization	7	5	3	0	0	2	0	1	2
Lexicalization	0	4	4	2	1	0	19	5	3
Omission	0	0	0	0	0	0	1	0	0
Unclassifiable	2	1	1	0	2	1	5	0	0
Total	22	20	14	8	4	4	43	18	6
Repetition	6	3			10		8		4
W > NW	**Y**						**Y**		Y
HF > LF	**Y**	**Y**			Y				**Y**
HI > LI									
Comprehension Def.	Anomic		Agnosic				None		

Table 4.5: Results For Patients With Evidence Of Reading Deficits
Only significantly low scores are reported.
Scores are standardized (z-scores)
Bold type marks deficits seen in both reading & repetition.

With the possible exception of AB (whose deficit is very mild) none of these patients neatly fits the pattern of pure (Type 1) surface dyslexia as defined by Shallice and McCarthy (1985): errors in reading irregular words only, with only regularization errors. Several patients (JH, ES, DM, and JD) fit the criteria for mixed (Type 2) surface dyslexia. Note, however, that the criteria for this syndrome (very poor reading of irregular words, with somewhat abnormal reading of regular and nonwords, and with a mixture of regularization and other errors) are very loose.

JH and ES both have speech output problems, which might account for their high rate of phonological errors. They also have the highest rate of regularization errors, making them the best examples of pure surface dyslexia. Their reading deficits are almost startlingly similar, down to the proportion (and, indeed, almost the number) of errors made.

As the result of their studies of PPA patients, Patterson and her colleagues have hypothesized that "the communication between semantic memory and the phonological output lexicon is partially responsible for holding the phonological elements of word together" (Patterson & Hodges, 1992), and, more precisely, that "deterioration of word meaning leads to a particular form of reading disorder designated...surface dyslexia", characterized by a difficulty reading orthographically irregular words. Although the data reported here relate directly to this hypothesis, there are two difficulties with interpreting the significance of the data. The first is that there are not enough patients in this study to test the hypothesis statistically. The second is that our reading words are not well-enough matched with the words in the comprehension tests, though the hypothesis is based on the idea that reading and semantic representation *of the same word* are related. Nevertheless, looking at the relation between scores on the naming tests and reading of irregular words gives us a rough measure of how well the hypothesis holds up. The assumptions are that the identified deficits in reading and naming will not be limited only to the specific exemplars in the tests, but will rather indicate a more general problem with semantic access or with reading, and that we may therefore expect to see some concordance of deficits even if the items on the two tests are not matched. The data collected for this study do not wholly support Patterson and Hodge's hypothesis. These data are graphed in Figure 4.13.

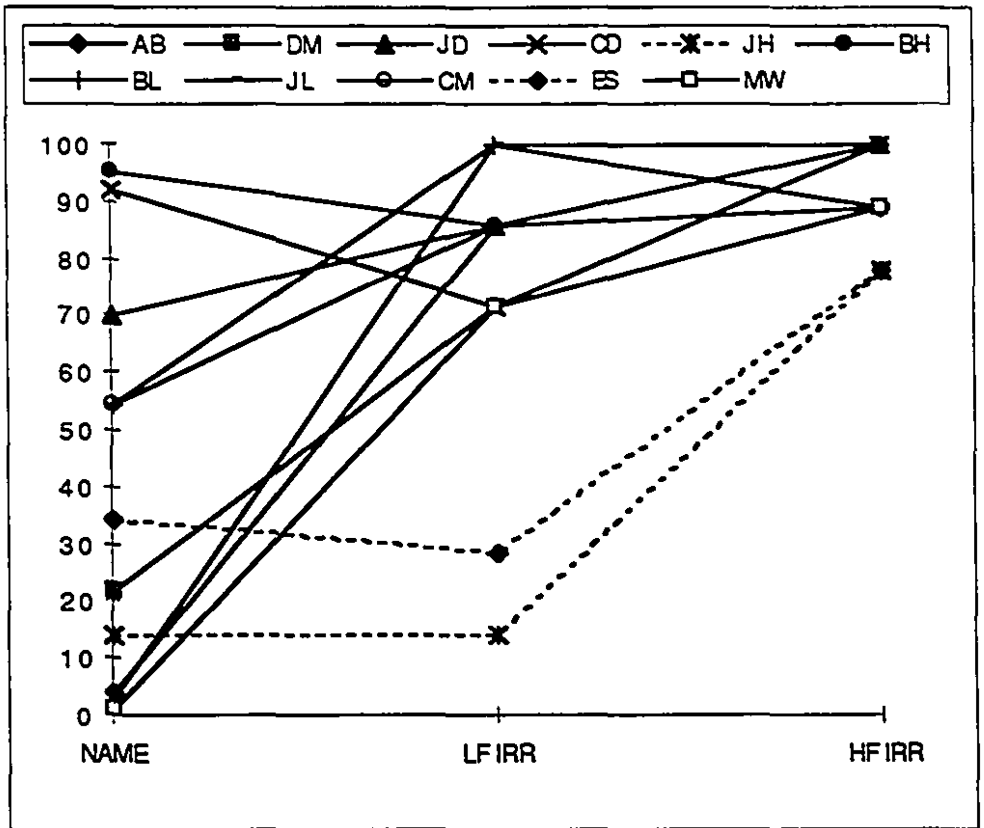


Figure 4.13
Naming deficits (averaged across modality) and reading of irregular words across all PPA patients.
(Because we do not have norms by frequency, data are charted as raw percents.)

It is immediately obvious that some patients (CM, DM, BL, MW, and, most notably, JL) who are extremely poor at naming and who have semantic deficits are easily able to read irregular words. JL can name only a single item (3%), but he read 100% of low frequency irregular words, and 89% of high frequency words. Similarly, BL correctly named only 54% of the items in the two naming tests, but was able to correctly read every irregular word she saw.

It is also clear that the only two patients with clear signs of surface dyslexia (ES and JH) are both among the worst at naming, and are by far the worst at reading irregular words in both frequency bands. This latter observation is not a necessary implication of their surface dyslexia, since the factors determining their surface dyslexia did not take into account absolute scores at reading irregular words, but only relative scores and their error pattern in reading.

These observations place the directionality of Patterson and Hodges' hypothesis into some doubt. It is not the case that a loss of word meaning is in itself sufficient to cause a concomitant deficit in reading irregular words. However, it may nevertheless be the case that loss of word meaning and surface dyslexia do tend to co-occur, either because they are both modulated by the same brain region or because there is some simple functional link between them. Patterson & Hodges (1992) do stress that the surface dyslexia and comprehension loss need not occur simultaneously, since the interaction between the semantic system and the lexical system can decay slowly.

Although it is particularly interesting that the two surface dyslexics in this study are also the only two clearly anomic (as opposed to agnosic) patients, it does not seem possible to draw any simple conclusions from this. Patterson and Hodges do not clearly distinguish between anomia and agnosia, using the term 'anomia' throughout their papers. Although some of their six patients (i.e. PP, PB) are clearly agnosic by my criteria, at least one (FM) is equally clearly (and others are apparently) anomic. Since some of their clearly agnosic patients showed the predicted effect, the relation between surface dyslexia and comprehension deficits can not be related in any simple fashion to the nature of the comprehension deficit, as the data from this study might have suggested.

Another unusual result in the reading data is the fact that, of the four patients who showed a significant dissociation between their ability to read abstract and concrete words, three showed an uncommon advantage for reading abstract words. (CM, who did not have a reading deficit, was also significantly better at reading abstract words than concrete words.) One of the three was DM, who has a well-documented abstract word sparing which goes beyond reading (see Appendix B for details). Although the remaining two patients (CD and BH) showed very small advantages for reading abstract words, they did not show low imagery advantages in any other test, and were both worse at abstract word comprehension than at concrete word comprehension. The data thus far are too weak to support any speculation about the sparing of abstract words as a symptom of PPA. However, this unusual reverse dissociation occurred often enough to imply that it may be worth attending to more closely in future research with PPAs.

Overall, the results of the reading tests are very much in keeping with results reviewed in Chapter 3, which found that reading is relatively intact in roughly 35% -50% of PPA patients, even quite late in the disease process. In this study, two patients had no reading deficits at all, and an additional four had only mild deficits. The remaining five (45.5%) had more severe deficits.

iv.a) Deficits in Affixed Word Comprehension

Perhaps the most unexpected finding in this study was the consistent difficulty across all eleven patients with comprehension of affixed words, a deficit which has not previously been associated with PPA in the literature . The affixed word comprehension deficits are summarized in Table 4.6, along with the analogous tests which do not rely on affixed word processing. All eleven patients scored low on at least two of the five tests of affixed word comprehension, and five patients were low on all five tests. In order for these low scores to be reasonably attributable to a deficiency in affixed word comprehension, it must be determined that the problem is not secondary to a more general deficit in word comprehension, by comparing performance on the analogous tests. The standardized results are graphed pair-wise in Figures 5.14 and 5.15.

	AB	DM	JD	CD	JH	BH	BL	JL	CM	ES	MW
ALD-DER	7	4	10	1	24	2	4	-	6	11	9
ALD	6	3	6	1	21	0	0	-	1	7	7
WLD-DER	6	5	10	1	7	2	2	7	2	6	11
WLD	6	15	3	0	8	2	2	5	2	4	15
A-WPM-DER	5	3	6	2	5	5	1	-	0	3	6
A-WPM	5	9	1	2	14	1	6	4	2	2	21
W-WPM-DER	3	1	3	7	4	2	1	-	2	5	3
W-WPM	0	6	1	1	4	0	0	5	2	1	10
W-SYN	3		6	3	5	6	4	-	5	4	3
Failed Tests	5	3	5	3	5	3	2	(1?)	2	5	5

Table 4.6: Affixed Word Comprehension (PPA Patients)
Standardized scores below average.
ALD/WLD = Auditory and written lexical decision of affixed words
A-WPM/W-WPM = Auditory and written lexical decision of affixed words
W-SYN - Written synonym judgments of affixed words

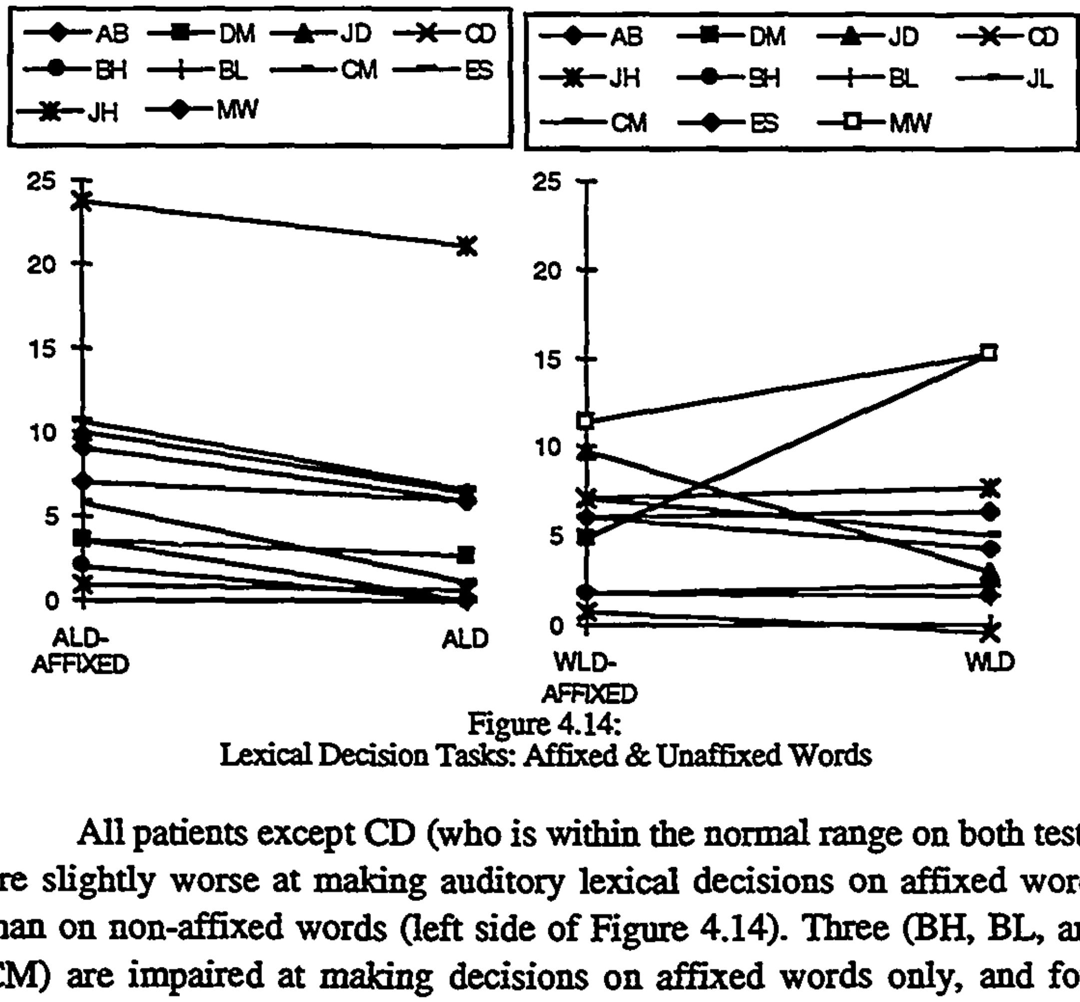


Figure 4.14:
Lexical Decision Tasks: Affixed & Unaffixed Words

All patients except CD (who is within the normal range on both tests) are slightly worse at making auditory lexical decisions on affixed words than on non-affixed words (left side of Figure 4.14). Three (BH, BL, and CM) are impaired at making decisions on affixed words only, and four others (JD, JH, ES, and MW) are three or more standard scores worse on affixed words than non-affixed words.

In the written modality most of these differences disappear (right side of Figure 4.14). Standardized performance on the two tests is remarkably similar for most patients, differing by two or fewer standard scores in eight cases. Two patients (MW and DM) are much worse (by four or more standard scores) at making written lexical decisions on non-affixed words than on affixed words. Only JD (seven standard scores difference) and ES (two standard scores difference) are worse at making written lexical decisions on affixed than non-affixed words

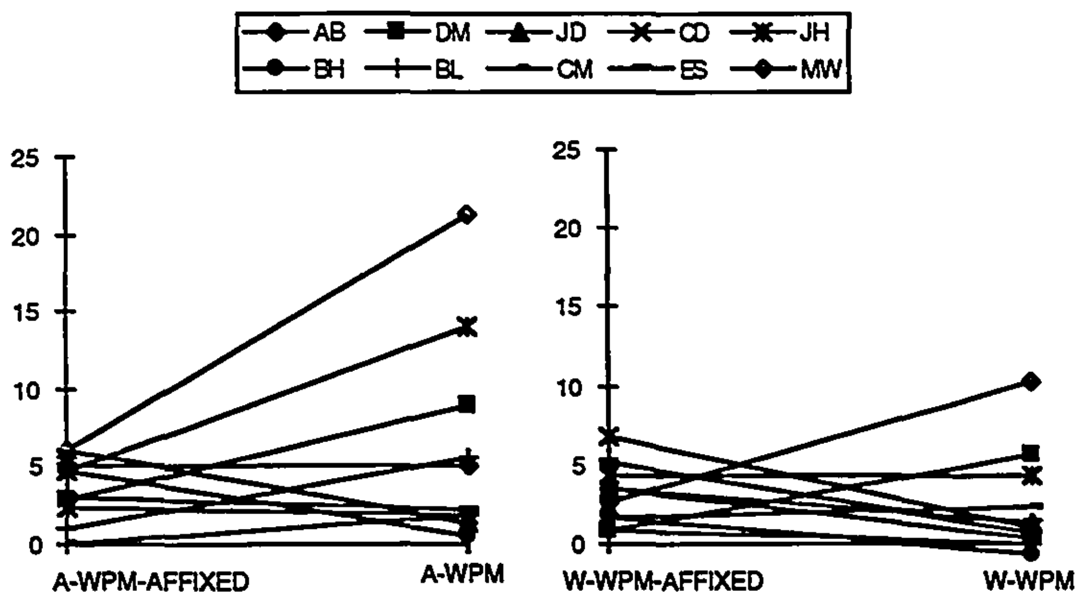


Figure 4.15: Word-Picture Matching Tasks
Affixed & Unaffixed Words

Four patients (JH, BL, DM, and MW) scored at least five standard scores better on the affixed auditory word picture matching test than on non-affixed version of the test (left side of Figure 4.15). One other (CM) scored two standard scores better on the affixed version. Three (AB, CD, and ES) scored approximately the same (fewer than two standard scores difference) on both tests. Two (JD and BH) of the remaining three scored at least two standard scores worse on the affixed word test than they did on the non-affixed word test. No comparison could be made for JL, who did not complete the affixed word version of the test.

The most notable result on the written word picture matching test comparisons (right side of Figure 4.15) is the massive advantage showed by MW (7 standard scores) and DM (5 standard scores) for matching affixed words to pictures over matching non-affixed words. Their large, consistent advantage in this direction on three of the four tests (and small difference on the remaining test) leaves little doubt that they have spared access to affixed words relative to non-affixed words. Such a result has never been documented in the literature, and, indeed, seems almost paradoxical. Since affixed words are clearly a more complex version of non-affixed words, one might no more expect to find a patient who can comprehend affixed but not unaffixed words than one might expect to find a circus performer who can juggle six balls, but not three. There are two possible explanations for the finding.

The first explanation, and perhaps the most intriguing explanation from a theoretical point of view, is that the patients who can understand affixed words better than unaffixed roots are able to glean some information about word type from understanding the meaning of the affix itself. This explanation makes a straight-forward prediction. If it is true, then the affixed word advantage should be less apparent in the lexical decision tests, since the foils in that test use legal affixes, making it impossible to gain any useful information from focusing only upon the affixes. It is not possible to test this post-hoc prediction statistically with only two relevant cases. However, we may get some measure of whether the predicted effect holds by examining the differences in standardized scores between affixed and non-affixed lexical decision for those patients who showed the most consistent affixed word advantage (see Figure 4.16). Both patients scored higher with affixed words on only a single one of the tests: auditory lexical decision. MW's scores on the affixed and nonaffixed lexical decision tests in both modalities are closer (by more than three standard scores) to each other than are her scores on the affixed and nonaffixed word-picture matching tests. These observations lend weak support to the hypothesis that the two patients are gaining information from the affixes. In DM's case, however, the overall pattern was not as hypothesized. Although the difference between his scores on the auditory lexical decision task was much smaller than his differences on any other pairs of tests, the difference in the written modality was the largest of the four differences. Overall the support for the hypothesis is quite weak.

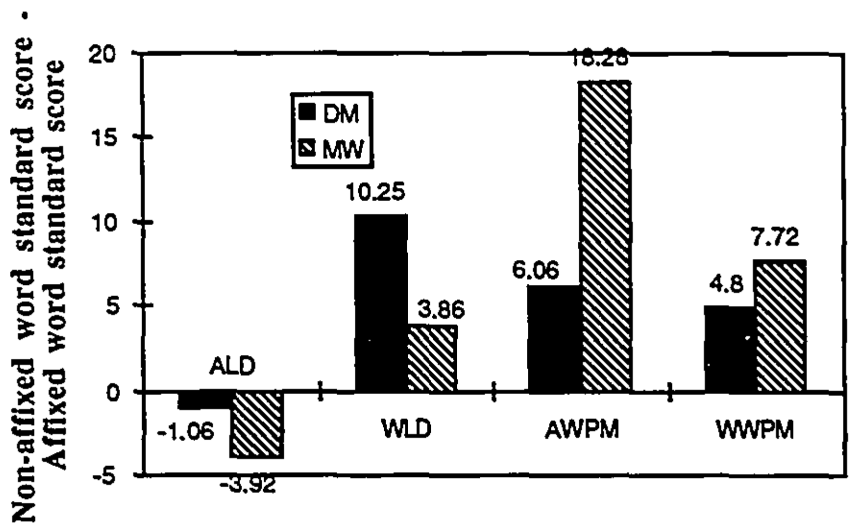


Figure 4.16
Differences between scores on analogous tests
of affixed and non-affixed word comprehension,
for patients with a consistent advantage for comprehension of affixed words

A second possible explanation for the phenomenon of affixed word sparing is simply that it is only the comprehension of nouns which has been impaired. Since many of the words in the tests of affixed word comprehension are not nouns, but all of the words in the tests of non-affixed word comprehension are nouns, a patient who is impaired only at nouns may be expected to achieve a higher score on the affixed word tests. Again we may test the plausibility of this prediction in an informal way, by breaking down the patient's performance on the affixed word tests by word type (see Figure 4.17). These data must be viewed as suggestive rather than conclusive, since the number of stimuli in each word type is too low (almost always less than 8) to allow us to draw any reliable statistical conclusions. Nevertheless, the trend is clear: nouns are by far the worst category for both patients on three of the four tests. This is supportive evidence in favour of this second explanation for the apparent phenomena of affixed word sparing.

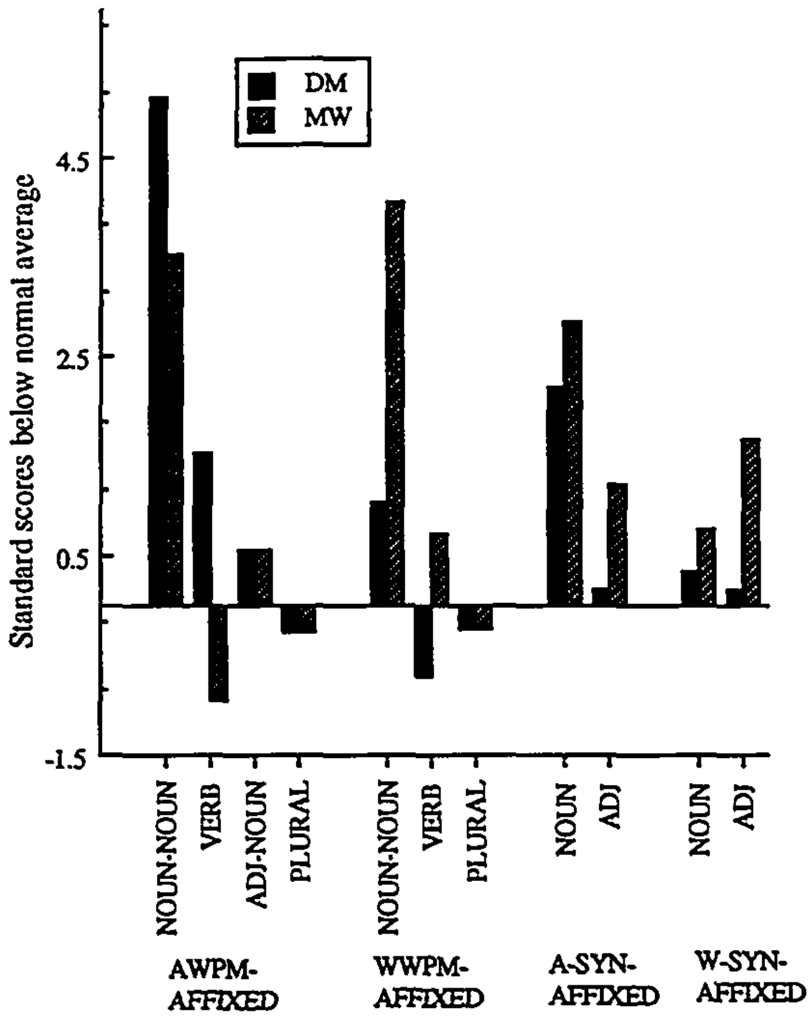


Figure 4.16
Performance on affixed word comprehension tests, by word subtype, for patients with a consistent advantage for comprehension of affixed words

If we can show that consistent differences are also possible in the opposite direction from DM and MW- that is, if we can find at least one patient who is clearly more impaired on affixed than non-affixed words- then we will have a double dissociation. Such a dissociation provides strong evidence that affixed word comprehension is functionally modularized (Shallice, 1988), though the discussion in the last two paragraphs suggests that this apparent modularization may prove to be an artifact. Such evidence is available. Among the five patients who scored at least two standard deviations better on the non-affixed version of the written word-picture matching test was JD, who thus shows a consistent advantage (by at least two standard scores, but by more than four standard scores in three of the four comparisons) for non-affixed words across all four pairs of comprehension tests. BH and ES showed a similar but smaller advantage on three of the four comparisons, suggesting that they too may have a comprehension deficit which is worse for affixed words than non-affixed words.

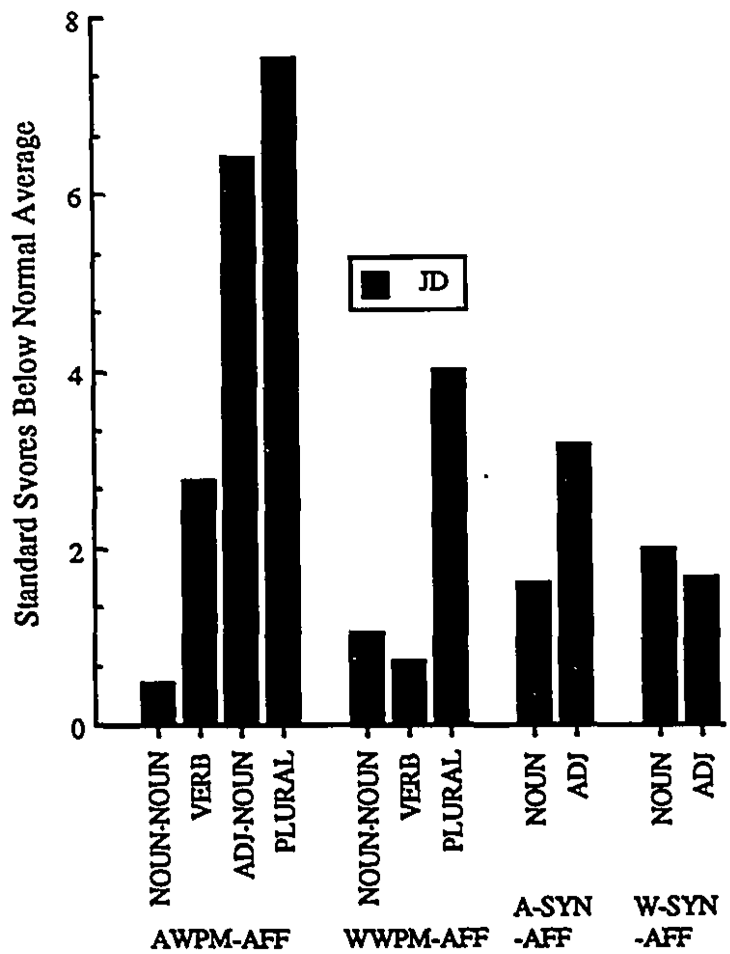


Figure 4.17
Performance on affixed word comprehension tests, by word subtype, for one patient with a consistent advantage for comprehension of non-affixed words

The identification of a single patient with a deficit which is clearly specific to non-affixed word comprehension allows a further test of the second hypothesis above. If the apparent affixed words sparing is indeed secondary to a more general impairment with the comprehension of nouns, then we should not expect to see a noun-specific impairment in patients with the inverse impairment. The relevant data is graphed in Figure 4.17 for the single patient with a clear non-affixed word comprehension deficit. Although it is again not possible to test the hypothesis statistically due to the low number of observations per cell, the trend is clearly in the direction predicted: nouns are not the worst category on three of the four tests, and are not much worse on the fourth test, a very different pattern from that of the two patients in Figure 4.16. This lends further suggestive support to the second explanation for the apparent phenomena of affixed word sparing: that patients who have an affixed word sparing show that pattern secondary to a noun-specific impairment. The interesting implication is that comprehension of affixed words may be modulated by information about word type.

It is not possible to test the inverse of the first hypothesis in a similar way, since the prediction to be tested is confounded with the way patients with a non-affixed word comprehension deficit are identified.

iv.b) Deficits in Affixed Word Production

	AB	DM	JD	CD	JH	BH	BL	JL	CM	ES	MW	Total
O-PROD	-	3	9	4	-		2	-	14	14	17	7
INF > DER			Y		-			-				1
L2 > L1		Y	Y		-			-	Y	Y		4
W-PROD	23	15	33	6	31	13	13	-	31	30	27	10
INF > DER	(Y)	Y		Y	Y	Y	Y	-	(Y)		Y	5
L2 > L1		Y				Y	Y	-			Y	4
True errors?	Y		Y	Y	Y					Y	Y	6

Table 4.7: Affixed Word Production
Standardized scores below average.
O-PROD/W-PROD = Oral and written production of affixed words
INF > DER = Significant advantage for inflected over derived words?
L2 > L1 = Significant advantage for Level 2 over Level 1 words?
Results in parentheses indicate a significant effect in the opposite direction to the label.

Although only one PPA patient (BH) scored within the normal range on the test of oral affixed word production, and none scored within the normal range on written affixed word production, just six of the eleven PPA patients (AB, JD, CD, JH, ES, and MW) were deemed on the basis of their error pattern to actually have a deficit which was specific to morphological production (see Table 4.7). All six were worse (and all except CD far worse) at written than spoken production.

No patient showed consistency between modalities with respect to the effect of the word type (inflectional versus derivational, or Level 1 versus Level 2) of the stimuli. This finding suggests that affixed word production may dissociate by modality.

Although two of the three patients who were worst at affixed word comprehension- JD and ES- were also among the worst at affixed word production, there is otherwise no apparent simple relation between patients who had difficulty with affixed word comprehension, and patients who had difficulty with affixed word production.

CONCLUSION

This chapter shows that the kinds of language processing deficits seen in PPA are widely variable. The most common deficit was a deficit in affixed word production, which was identified, at least tentatively, in all eleven subjects. Such a deficit has not been previously described as a diagnostic sign of PPA. Reading and naming deficits were the next most common deficits, each identified in eight patients. The nature of these deficits was quite variable. Deficits in sentence production and comprehension and in affixed word comprehension were also common.

In the next chapter, results from the DAT patients are presented and analyzed to investigate whether these DAT patients exhibit a similar variability. In chapter 6, the two groups are compared to see if there are any features which allow patients from the two groups to be clearly distinguished.

Chapter Five: DAT Analysis

Trace, re-trace, tide-worn wash of mind.
There is nothing left to strip away, grind
Down, wear off: but still not pure enough, no
Clarity. Words stumble, clutter, clog. I remain
A draughtsman; thought, dull pencil used
To trace the outlines that fragment and blur
At every stroke.

Jan Zwicky
Wittgenstein Elegies

Language deficits have been recognized to be a symptom of senile dementia of Alzheimer's type (DAT) since 1907, when Alois Alzheimer described the first patient with the disease that now bears his name (see Mathews, Obler, and Albert, 1994, for a detailed review of that case). Recent studies have confirmed that language deficits are almost always seen in DAT patients. Appell et al (1982) found language deficits in all 25 patients in their study, Emery and Emery (1983) found deficits in every one of 18 DAT patients, Cummings et al (1985) found deficits in each of 30 DAT patients they looked at, and Faber-Langendoen et al (1988) documented language problems in 24 of 66 (36%) of their mild DAT patients and all 17 (100%) of their severe DAT patients.

Despite this long history and the proliferation of studies confirming the ubiquity of language deficits in DAT patients, however, there have been no broad, quantitative studies designed specifically to elucidate in detail the possible ways that language can deteriorate in DAT patients. No published study of language deterioration in DAT patients has allowed for a systematic comparison of the pattern of language deterioration between individual DAT patients assessed with the same instrument, since all published studies with more than a single subject have masked individual differences by reporting only group means. This is particularly problematic in light of two results from Faber-Langendoen et al's (1988) paper. One problematic result is the implication that aphasic patients may constitute a subgroup of DAT patients, rather than reflecting the general cognitive decline which is a defining characteristic of the disease- so that the group results may be averaging across qualitatively different subgroups. The second problematic result, for a similar reason, is the finding in the same paper that all language deficits are highly correlated at the group level. No published study has consistently reported possible cross-test effects such as word frequency effects, rendering theoretical interpretation of the results problematic. This lack of experimental control reflects the fact that studies published to date have often tended to rely upon impoverished stimulus sets, which do not contain enough stimuli to allow for thorough, interpretable testing of each identified linguistic subfunction with appropriate control for frequency, imageability, and other word characteristics which may affect performance. Most studies to date have also failed to differentiate precisely enough between different levels of linguistic processing (e.g. morphological, phonological, syntactic, and semantic levels). Although my main purpose in this chapter is to compare the performance of DAT patients to the performance of PPA patients, the data presented in this chapter is also an attempt to remedy the failings of previous work by reporting individual results, interpreted in a cognitive neuropsychological framework, of a detailed, systematic, hierarchically-structured study of language deficits in eleven DAT patients.

LANGUAGE DEFICITS IN DAT: A REVIEW

Despite the fact that the pattern of symptoms among individuals is masked in previous studies, there is some agreement on the kinds of language deficits and spared functions which are most commonly associated with DAT.

Among the high-level functions, there is widespread agreement that syntax is spared (Whitaker, 1976; Schwartz et al, 1979; Appell et al, 1982; Murdoch and Chenery, 1987; Patel & Staz, 1992), although this sparing is sometimes obscured by deficits in the use of phrase markers, failure to make grammatical agreements, and by a tendency of DAT patients to leave phrases unfinished (Constantinidis et al, 1978). Sentence level comprehension deficits, especially of abstract or complex propositions, are commonly reported (Appell et al, 1982; Emery and Emery, 1983; Cummings et al, 1985; Murdoch & Chenery, 1987). A well-documented deficit at the level of semantic processing (Irigaray, 1967; Constantinidis et al, 1978; Schwartz et al, 1979; Martin & Fedio, 1983; Warrington, 1975) appears to be related to a deficit in accessing information about special attributes (within-category specifiers) and subordinate category information, while superordinate categorical information is usually left intact (Flicker et al, 1987). Appel et al (1982) have aptly described this deficit as a "loss of differentiation in the semantic field".

The most commonly reported low-level deficit (Sim and Sussman, 1962; Irigaray, 1967; Barker & Lawson, 1968; Appell et al, 1982; Bayles & Tomoeda, 1983; Rosen, 1983; Kirshner et al, 1984; Cummings et al, 1985; Hier et al, 1985; Flicker, 1987; Murdoch & Chenery, 1987) is word-finding difficulty, with an associated anomia, which may be modulated by word frequency (Appell et al, 1982). The word-finding deficit has been shown to be worse in aphasic DAT patients than in similarly-demented but non-aphasic patients (Faber-Langendoen et al, 1988), indicating that the impairment is not due only to the fact that word-generation tests tap cognitive processes other than language functions. Some DAT patients show impaired naming as their only language function deficit (Kirshner et al, 1984). Flicker et al (1987) have noted that a decline in naming ability is of particular interest to the diagnostician, since it is one of the few language deficits which is almost never seen as an effect of normal aging.

Semantic paraphasias have been frequently reported among DAT patients (Constantanidis et al, 1978; Schwartz et al, 1979; Appel et al, 1982; Bayles & Tomoeda, 1983; Martin & Fedio, 1983; Murdoch & Chenery, 1987). Phonological paraphasias are sometimes seen in the later stages of the disease (Constantanidis et al, 1978; Horner & Heyman, 1982). Repetition (Beattie & Emery, 1983; Cummings et al, 1985; Murdoch & Chenery, 1987), reading aloud (Benson, 1982; Cummings, 1985; Cummings, 1986; Patel & Staz, 1992), and phonological discrimination (Schwartz et al, 1979; Cummings, 1986; Murdoch and Chenery, 1987) are usually spared relative to the other low-level functions described, especially in the early stages of the disease.

GLOBAL OVERVIEW

The case studies of the eleven DAT patients are presented in Appendix C. In this section the deficits identified in that appendix will be summarized and analyzed. The identified deficits are presented in Table 5.1.

Three patients (AB, EF, and DO) failed or were unable to complete nearly every test in the PAL battery. Because it is not possible in these cases to distinguish between tests failed for reasons which are secondary to the language system (general dementia) and tests failed due specifically to damage to the language system, I have, with one exception, not attempted to code the deficits identified in those patients. The exception is EF's category specificity in naming, which may be flagged with some confidence since it was not a deficit but a category-specific sparing, of the fruits and vegetables category. Despite the difficulty in analyzing results for these patients, two similarities are worth noting because they point to relative preservation of function. The first is that all three patients made a great many semantic errors in naming in both modalities, suggesting that they had inaccurate, as opposed to totally obliterated, access to semantics. This is consistent with Appel et al's (1982) description of the semantic deficit in DAT "loss of differentiation in the semantic field". The similarity among the three patients is that all three had relatively preserved repetition (especially relative to reading), with the bulk of their repetition errors being phonological paraphasias, an observation which is in accordance with the summary of language deficits in DAT provided at the beginning of this chapter.

Subjects	NB	DD	MR	DS	JS	RS	YS	Total	AB	EF	DO
Auditory Input deficit	Y				(Y)	(N)		2	-	-	-
Readng Deficit		Y			Y			2	-	-	-
Anomia		Y	(W)	Y				3	-	-	-
Agnosia	Y					Y	Y	3	-	-	-
Objects < Organic							(Y)	1	-	-	-
Organic < Objects								0	-	(Y)	-
Bad Abstract Words			A					1	-	-	-
Abstract word sparing								0	-	-	-
Morphology Deficits											
Production		(Y)	W		(Y)	(Y)	(Y)	5	-	-	-
Comprehension		(Y)	W		(Y)		(Y)	4	-	-	-
Sentence Deficits											
Production									-	-	-
Comprehension	W	Y	A					3	-	-	-

Table 5.1: Summary Of DAT Deficits Identified With PAL
Y = Deficit is present; N (or blank) = Deficit is absent; - = Deficit is unidentifiable
Deficits in parenthesis are equivocal.
R = Reversible sentences only (Syntax Deficit)
A = Auditory Modality Only; W = Written Modality Only

Because the severity of their deficits renders questionable any conclusions we might draw about the language system, I will remove AB, EF, and DO from further consideration here.

The remaining eight DAT patients have a small number of the syndromes (listed in Table 5.2) identified among the PPA patients. The average number was 3.6. However, since many of those are equivocal (as Tabie 5.1 makes clear), this number represents an upper bound which may err on the high side.

Although the low number of taxonomic characters (syndromes) renders the measures of dissimilarity less useful than they might otherwise be, since there are only small differences between patients, a multi-dimensional scaling using a linear Kruskal loss function was still possible. The Shephard diagram was smooth and a stable two dimensional configuration, accounting for 99.9% of the variance, was found after 49 iterations. The resultant scaling is reproduced in Figure 5.1.

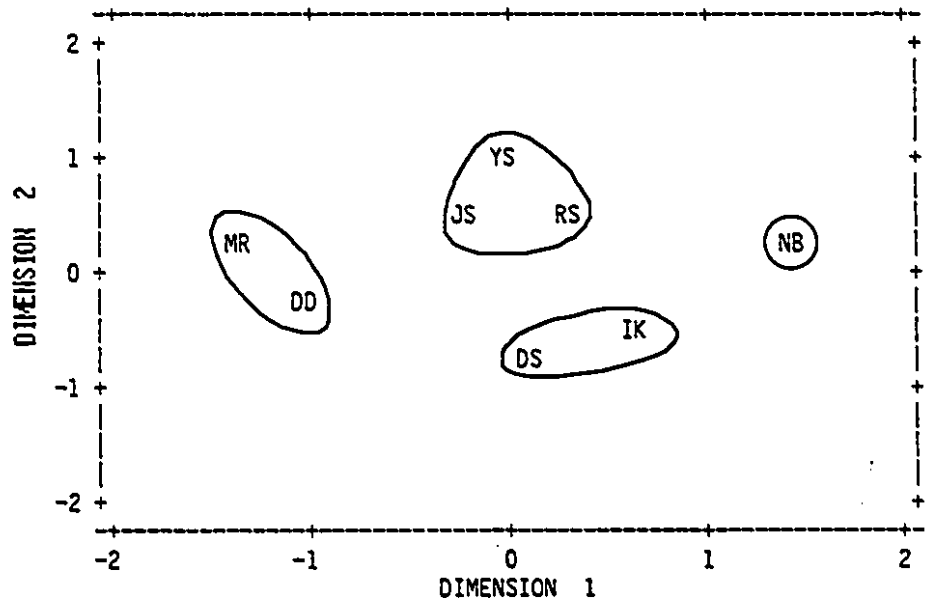


Figure 5.1
Multi-Dimensional Scaling Of DAT Patients
By Euclidean Distance In 'Symptom Space'
(see Table 5.1 and accompanying text)

The multi-dimensional scaling does reveal some weak regularities in the distribution of the syndromes, with four weakly-defined clusters identifiable. It is important to be stress that multi-dimensional scaling is an atheoretical technique with respect to psycholinguistics. The resultant clusterings may be useful as a starting point for theorizing, but they must not be taken to instantiate theories in themselves. The data are reorganized according to the clusters in Table 5.2.

Cluster #	1		2			3		4
Subjects	MR	DD	YS	JS	RS	IK	DS	NB
Auditory Input deficit				(Y)	(N)			Y
Reading Deficit		Y						
Anomia	(W)	Y					Y	
Agnosia			Y		Y			Y
Objects < Organic			(Y)					
Organic < Objects								
Bad Abstract Words	A							
Abstract word sparing								
Morphology Deficits								
Production	W	(Y)	(Y)	(Y)	(Y)			
Comprehension	W	(Y)	(Y)	(Y)				
Sentence Deficits								
Production						Y		
Comprehension	A	Y						W

Table 5.1: Summary Of DAT Deficits Identified With PAL
Y = Deficit is present; N (or blank) = Deficit is absent
Deficits in parenthesis are equivocal.
A = Auditory Modality Only; W = Written Modality Only

Cluster 1

The first cluster contains two patients, MR and DD. The positive similarities between these two patients are that they both showed evidence of an anomia, a deficit in morphological production and comprehension, and a deficit in sentence comprehension. There is no known psycholinguistic coherence between these symptoms.

Cluster 2

The second cluster consists of YS, JS, and RS. They were similar to the patients in cluster 1 inasmuch as all three showed evidence of a deficit in morphological production, and two of the three also showed evidence of a deficit in morphological comprehension. YS and RS were both agnosic.

Cluster 3

Cluster 3 consists of two patients who are similar only insofar as they each have just one deficit, and in both cases the deficit is not common among the other patients. IK had a sentence production deficit, and DS was anomic.

Patient NB

NB was not clustered, because he has three deficits (an auditory input deficit, an agnosia, and a deficit in written sentence comprehension) which are not seen together in any other patient.

ANALYSIS BY DEFICIT

In order to facilitate the comparison of results between the two diagnostic groups in the next chapter, I will briefly consider the performance of the DAT patients on the four sets of deficits which were considered in more detail among the PPA patients: the auditory input deficits, the naming deficits, the reading deficits, and the deficits in affixed word production and comprehension.

i.) Auditory Comprehension Deficits

Only a single patient (NB) showed evidence of a likely (albeit very mild) auditory input deficit, based upon his poor scores in phoneme discrimination and repetition, especially of low frequency words and nonwords. However, he did not show any consistent difficulty with auditory input on any of the tests of comprehension except for a mild difficulty with auditory (but not written) word-picture matching.

JS's deficit is more equivocal. He had poor repetition (again of low frequency words and nonwords) and a score on the phoneme discrimination test that was borderline (1.8 SDs below normal). Like NB, he did not show any consistent difficulty with auditory input on any of the tests of comprehension.

ii.) Naming Deficits

All but two (DS and IK) of the DAT patients showed evidence of either an anomia or an agnosia. Although in MR's case the evidence for an anomia is highly equivocal, she has been included because she (unlike IK) would have scored significantly low on the written naming test even if her orthographic errors were not counted as erroneous. The relevant results are summarized in Table 5.2.

	Anomia			Agnosia		
Patient	DD	MR	DS	NB	RS	YS
AWP	2		3	3	6	4
WWP			2		3	3
O-NAME	13		12	6	5	8
W-NAME	14	6	22	17	17	31
SEM.				3	4	3
Category Specificity?						
AWP		FV>AN FV>OB OB>AN	FV>AN OB>AN	OB>AN FV>AN	ORG>OB FV>AN FV>OB	FV>AN FV>OB OB>AN
O-NAME	OB>ORG OB>AN OB>FV		ORG>OB FV>AN FV>OB		OB>AN FV>AN	FV>AN FV>OB OB>AN
WWP	OB>ORG	IN>ORG	OB>ORG	OB>ORG OB>FV	FV>OB	FV>AN
W-NAME	OB>ORG AN>FV OB>FV	AN>FV OB>AN	AN>FV AN>OB	AN>FV OB>FV	ORG>OB AN>OB FV>OB	

Table 5.2
Results For Patients With Evidence Of Naming Deficits
Only significantly low scores are reported.
AWP = Auditory word-picture matching
WWP = Written word-picture matching
O-NAME = Oral picture naming
W-NAME = Written picture naming
SEM. = Global semantics battery score
Numbers are standard scores below average.
Reported dissociations use a 2 SD cut-off point.

DD and DS are both clearly anomic, with very low scores on both tests of naming, but no deficit on the semantics battery. DD's deficit seems clearly to be worse for organic than inorganic stimuli, and is particularly apparent with the naming of fruits and vegetables. Although DS's pattern is less consistent, there is some evidence of a particular deficit in naming inorganic objects.

NB, RS, and YS are all agnosic, scoring significantly low on both the naming tests and the semantics battery. Although all three showed significant dissociations by category, there is not enough consistency in the distribution of these effects within any single patient to allow us to ascribe any of them a category-specific deficit.

iii.) Reading Deficits

Despite the high normal standards on the reading test, only one (DD) of the eight DAT patients being considered scored significantly (albeit mildly) low at reading (see Table 5.3). His mistakes were split between short irregular words and short orthographically complex nonwords. The deficit is too mild to interpret in detail.

	DD
Reading	4
W > NW	
HF > LF	
HI > LI	
S > L	(Y)
REG > IRR	Y
PSIM > PCOM	
OSIM > OCOM	Y
Error Counts:	
Semantic	0
Phonological	2
Regularization	2
Lexicalization	0
Omission	0
Unclassifiable	0
Total	4
Repetition	
W > NW	
HF > LF	
HI > LI	
Comprehension Def.	Anomia

Table 5.3: Summary of the sole reading deficit identified among the DAT subjects

iv.a) Deficits in Affixed Word Comprehension

It is interesting to note, in light of the ubiquity of morphological processing deficits among the PPA patients, that such deficits are almost nonexistent among the DAT patients. Although five DAT patients (MR, DD, YS, RS, and JS) scored significantly low on one or more tests of affixed word comprehension or production, none of them could be unambiguously deemed on the basis of the nature of their productive errors to have a deficit that was clearly specific to morphological processing.

The relevant standardized results for the tests of affixed word comprehension are summarized in Table 5.4.

Subjects	DD	MR	JS	RS	YS
ALD-AFFIXED	3	0	3	3	5
ALD	2	0	3	1	0
WLD-AFFIXED	3	1	3	4	3
WLD	2	0	1	0	1
A-WPM-AFFIXED	1	1	1	0	1
A-WPM	2	1	0	6	4
W-WPM-AFFIXED	4	5	0	0	0
W-WPM	1	0	0	3	3
W-SYN	1	2	0	0	1
Failed Tests	3	2	2	2	2

Table 5.4: Affixed Word Comprehension (DAT Patients)
Standardized scores below average.
ALD/WLD = Auditory and written lexical decision of affixed words
A-WPM/W-WPM = Auditory and written lexical decision of affixed words
W-SYN - Written synonym judgments of affixed words

DD was the only patient who scored significantly low on more than two of the five tests of affixed word processing. However, his scores on the tests of lexical decision of affixed words were only one standard deviation below his scores for the general tests of lexical decision, making it unlikely that his low scores on the affixed word test indicate a difficulty specific to affixed words.

MR scored significantly low only on the two tests of written comprehension of affixed words which are semantically-mediated, word picture matching and synonym judgment. This may reflect a modality-specific deficit for affixed word processing, since (as explained below) MR also failed on the test of affixed word production in the written modality only.

The remaining three patients who scored low on tests of affixed word comprehension (JS, RS, and YS) all scored low only on the two lexical decision tests, the nonsemantically mediated tests of affixed word comprehension. Only JS showed an analogous deficit on the tests of non-affixed word lexical decision, and only in the auditory modality. These three patients may have a true deficit in comprehension of affixed words when they cannot constrain the comprehension with semantic information.

Both RS and YS were much better (by at least 3 standard scores) at matching affixed words to pictures than they were at matching unaffixed words to pictures. These were the only cases which showed dissociations of two or more standard scores in favour of affixed word comprehension over unaffixed word comprehension.

iv.b) Deficits in Affixed Word Production

The same five patients who were identified as showing possible indications of a deficit in affixed word comprehension also showed deficits in affixed word production (see Table 5.5). No patient made more than one error which was clearly a true affixation error, and only JS, RS, and MR made even one such error.

	MR	DD	YS	JS	RS	Total
O-PROD	3	3	6	2	4	5
INF > DER			(Y)			1
L2 > L1	Y	Y	Y	Y	(Y)	5
W-PROD	-	3	20	1	4	3
INF > DER	-	Y	Y			2
L2 > L1	-	Y	Y			2
True errors?	Y	N	N	Y	Y	3

Table 5.5: Affixed Word Production
Standardized scores below average.
O-PROD/W-PROD = Oral and written production of affixed words
INF > DER = Significant advantage for inflected over derived words?
L2 > L1 = Significant advantage for Level 2 over Level 1 words?
Results in parentheses indicate a significant effect in the opposite direction to the label.

Like the PPA patients, no patient showed consistency between modalities with respect to the effect of the word type (inflectional versus derivational, or level 1 versus level 2) of the stimuli. Unlike the PPA patients, the DAT patients did not score consistently lower at written production of affixed words than oral production.

CONCLUSION

The DAT patients may be divided into two groups with respect to their performance on the PAL battery. The three most severely demented patients were almost totally unable to complete any of the tests, for reasons which are probably not specific to their language deficits. The remaining eight DAT patients, in contrast, had very few language deficits. The most common clearly-defined deficit was an agnosia, which was the only one of the syndromes examined here which was seen unambiguously in as many as three of the eight testable DAT patients. Mild deficits in processing word morphology were seen more often, but in terms that were quite ambiguous. Sentence comprehension and production was remarkably intact among the DAT patients. There was no consistency among the three most mildly affected patients (each with two or fewer syndromes) so it is not possible to draw any conclusions about which aspects of language processing are likely to be the first affected by the disease process.

Chapter Six: Comparisons

The road to wisdom?
Well, it's plain and simple to express:
Err and err and err again
but less and less and less.

Piet Hein
As quoted in: Dan Dennett
Darwin's Dangerous Idea

In this chapter, the results from the previous two chapters will be compared and contrasted.

HYPOTHESES REVISITED

In the first chapter, I presented five neurologically-grounded hypotheses on the kinds of differences that might be expected between the language deficits in PPA and DAT. We are now in a position to move from an individual to a group perspective in order to assess those hypotheses. Because the first hypotheses is a general one, it will not be evaluated until all other hypotheses have been considered. I therefore begin here with an assessment of the second hypothesis.

Hypothesis 2

Hypothesis 2 was based on the differences between deficits associated with the posterior superior temporal lobe (Brodmann's area 40), likely to be affected in both PPA and DAT, and the mid-temporal lobe (Brodmann's area 21), which was judged to be more likely to be affected in DAT than in PPA. The hypothesis was stated in the following manner:

> DAT patients will show greater deficits than PPA patients in tasks known to be modulated through area 21: accessing names in the picture-naming tasks; verbal comprehension; repetition; reading irregular words; and writing. Both DAT and PPA patients will show deficits on tasks known to be modulated through area 40: speech production; phoneme discrimination; and reading nonwords.

It is immediately obvious, in light of the huge number of language problems documented among the PPA patients in Chapter 4, and the few documented in Chapter 6 among the DAT patients, that this hypothesis is not supported by the data. In fact, as Table 6.1 and Figure 6.1 make clear, PPA patients showed much larger deficits in every task mentioned in the hypothesis.

	PPA				DAT			
	% < 2 SDs	Ave.	SD	N	% < 2 SDs	Ave.	SD	N
AWP	64	6.1	6.4	11	63	2.4	2.2	8
A-COMP-ABS	100	5.8	3.0	8	25	1.8	2.2	8
A-SYN-AFF	50	2.6	2.5	8	0	0.6	0.6	8
A-MATCH-AFF	90	3.3	1.6	10	13	0.6	0.9	8
READING IRR	64	9.6	9.7	11	43	1.9	2.9	7
O-NAME	73	8.9	7.5	11	63	4.6	4.2	8
W-NAME	90	25.0	18.4	10	88	13.9	10.2	8
PD	70	3.8	3.9	10	13	1.5	1.3	8
READING NW	73	10.0	11.0	11	14	0.2	1.2	7
Orthographic Error Counts	% > 0	Ave.	SD	N	% > 0	Ave.	SD	N
Naming	80	3	2.8	10	63	1.8	2.4	8

Table 6.1
Group comparison of DAT and PPA patients on the tests relevant to Hypothesis 2 (see text).
All averages and standard deviations are standard scores except for the orthographic errors count.

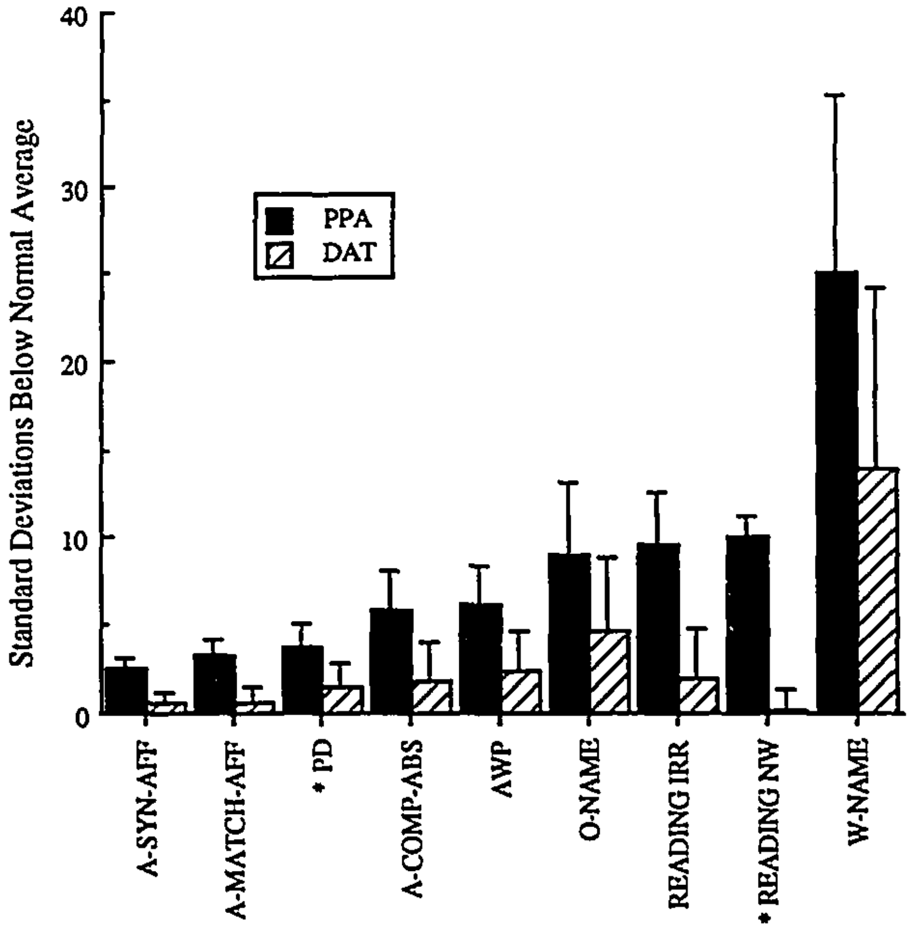


Figure 6.1
Group comparison of DAT and PPA patients on the tests relevant to Hypothesis 2 (see text).
Error bars are standard deviations.
Tests marked with an asterisk were those hypothesized to be equally affected.

We may easily dismiss the second hypothesis as not supported even weakly by the data presented here.

Hypothesis 3

The third hypothesis was based on expected differences in the involvement of frontal areas in the two disease processes under consideration. I hypothesized that:

> PPA patients will show greater deficits than DAT patients in tasks known to be modulated through Broca's area: tasks involving oral production of language, and sentence comprehension, especially of semantically reversible sentences.

The most important tests for this hypothesis were the repetition test and the sentence comprehension tests. We may also examine the number of phonological errors made in reading and oral naming. The relevant data are summarized in Table 6.2.

	PPA				DAT			
	% < 2 SDs	Ave.	SD	N	% < 2 SDs	Ave.	SD	N
Repetition	55	3.3	4.1	11	38	2	3	8
Oral Sentence Comprehension								
Reversible	63	2.6	1.4	8	38	1.1	2	8
Constrained	38	2.0	2.0	8	0	0.6	0.7	8
Written Sentence Comprehension								
Reversible	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Constrained	45	3.2	4	11	43	2.2	0.9	7
Phonological Error Counts	% > 0	Ave.	SD	N	% > 0	Ave.	SD	N
Reading	100	6.2	6.2	11	38	2.6	5.1	8
Naming	55	1.5	1.8	11	0	0	0	8

Table 6.2
Group comparison of DAT and PPA patients on the tests relevant to Hypothesis 3 (see text).

The scores on the repetition test do not lend support to the hypothesis. Six (55%) of the eleven PPA patients scored significantly low on the repetition task, compared to 3 of 8 (38%) DAT patients who scored low. This difference in proportion is not significant (Fisher's Exact Probability = 0.49). Moreover, there was no significant difference between the two groups ($t(17.0) = 0.75$; $p > 0.05$) on overall scores on the repetition task.

The repetition task is a weak test of oral production, since repetition may be impaired for a number of reasons. Slightly stronger evidence of a deficit specific to oral production comes from examining the incidence of phonological errors on the tests of reading and oral naming. All eleven

PPA patients made at least one phonological error in reading, and 55% (6 of 11) made at least one such error in oral naming. In contrast, only 38% (3 of 8) DAT patients made at least one such error in reading, and none made a phonological error in naming. These differences in ratio are not significant (Reading: Fisher's Exact Probability = 0.19; Naming: Fisher's Exact Probability = 0.07). PPA patients averaged 6.2 phonological errors in reading, with six patients making five or more such errors. The average number of phonological errors in reading made by DAT patients was 2.6, but only a single patient made more than two such errors. The analogous counts from the test of oral naming are not as dramatic in size, but only PPA patients made any phonological errors in oral naming.

In order to consider whether the differences in the number of errors are statistically significant, we must consider carefully what it is we want to measure. Clearly, it is not the simply the absolute number of phonological errors made by each subject in each group, since this measure will tell us nothing about phonological errors per se if one group simply makes more errors of all kinds. We are interested rather in the relative number of errors made by each individual which were phonological. However, merely stating each individual's phonological errors as a proportion of their own total errors is unsatisfactory for another reason, which is that it gives undue and unjustifiable weight to patients who made a small number of total errors. The most appropriate way of analyzing such data is to weight the contribution of each individual's phonological errors to the total errors made on that test within that group. This is easily achieved by dividing each individual's phonological error count by the total number of errors of all kinds made by that person's diagnostic group as a whole. The difference between the groups in the contribution made by phonological errors to the total errors was not significant ($t(7.9) = -0.47$, $P > 0.05$) for the reading test. It was not possible to calculate the statistic for the naming test, since there were no non-zero data points in the DAT group.

PPA patients scored lower on average in all sentence comprehension tests. A slightly greater number of PPA patients (5 of the 8 who completed the test) scored below norms on the reversible sentences than DAT patients (3 of 8), although the difference was not significant (Fisher's Exact Probability = 0.44). However, when the data are examined by individual (Table 6.3) it is clear only one PPA patient (but none of the DAT patients) showed an unambiguous dissociation (of at least 2 standard scores) in the predicted direction of being worse at reversible sentences than constrained sentences.

PPA Patients	AB	DM	JD	CD	JH	BH	BL	JL	AC	ES	MW
Repetition	9.8	-0.7	8.4	4.1	5.8	-0.2	-0.7	7.9	-0.4	3.4	-0.9
Oral Sentence Comprehension											
Reversible	-	0.6	1.3	2.8	4.4	-	-	-	4.4	2.8	2.1
Constrained	-	2.4	0.6	1.5	6	-	-	-	- 0	1.5	2.4
Phonological Error Counts											
Reading	1	5	18	0	13	12	6	1	1	10	1
Naming	1	7	11	0	0	2	1	1	2	8	0
Total Symptoms	3	2	4	1	2	3	2	3	2	3	1

DAT Patients	NB	DD	IK	MR	DS	JS	RS	YS
Repetition	4.1	-0.2	-0.9	-0.4	0.8	5.3	7.2	0.8
Oral Sentence Comprehension								
Reversible	- 0	5.1	1.3	2.1	0.6	- 2	- 0	2.1
Constrained	0.6	0.6	0.6	1.5	- 0	- 0	1.5	0.6
Phonological Error Counts								
Reading	0?	2	0	0	0	14	0	2
Naming	0	0	0	0	0	0	0	0
Total Symptoms	1	1	0	0	0	2	1	1

Table 6.3
Data from hypothesis 3 by individual patient
Bold text marks a data point predicted by the hypothesis.

It is easier to judge how well the hypothesis was supported by examining at the individual data. All the PPA subjects had at least one of the predicted symptoms, only two PPA patients had just one, and five had three or more of the symptoms. In contrast, three DAT patients had none of the symptoms, four of the eight DAT patients had only one, and none had more than two. The discontinuity of error types in naming is particularly noteworthy since it is not merely a reflection of a difference in ability to name: although DAT patients scored, on average, about twice as high as PPA patients, the DAT group as a whole still scored well below norms in oral naming. Evidently, the DAT patients were making errors in oral naming, but not phonological errors. I will return to this point below, when I discuss the difference in semantic errors in naming.

Additional support for the original formulation of this hypothesis comes from the observation that at least two (JH and ES) of the PPA patients (and possibly three, if we include JL) showed the 'classic' defining symptom of Broca's aphasia, a disturbance of expressive language which renders the patient nearly mute. (However, note that of these three, only JL is clearly worse at comprehension of constrained sentences.) None of the DAT patients showed this symptom clearly.

Although there is some suggestive support at the individual level for this hypothesis, it must be acknowledged that much of the support may be an artifact , since PPA patients generally have more deficits of all kinds. Any hypothesis which predicted more pathology in that direction would be likely to be supported. The fact that neither of the two clearest Broca's aphasics showed all four of the predicted symptoms casts some doubt on the assumptions behind this hypothesis.

Hypothesis 4

The fourth hypothesis was stated as follows:

> Because they are more likely than DAT patients to have neuronal degeneration extending into the parietal lobe to the junction of the posterior angular gyrus, and the parietal-occipital border, PPA patients will make more orthographic errors in tests of writing than DAT patients.

The data relevant to this hypothesis have already been presented in Table 8.1 above. Eight of ten PPA patients (80%) and five of eight (63%) DAT patients made at least one orthographic error on the test of written naming. These proportions are not significantly different (Fisher's Exact Probability = 0.51).

Using the same measure of weighted contribution to total errors as was described in the last section, we may be assured that there is in fact no significant difference between the number of orthographic errors made in the two groups ($t(11.0) = 0.59$; $P > 0.05$), and thus may confidently reject this hypothesis.

Hypothesis 5

The last hypothesis concerned the nature of semantic impairment in each of the groups. In order to make it easier to assess, it is re-stated here as a pair of conjunctive statements:

> i.) DAT patients will be more impaired on the semantic battery tests with both written and picture stimuli than PPA patients.
> ii.) PPA patients will show more difficulty in naming pictures, especially in the oral modality than DAT patients.

As shown in Figure 6.2, group average scores and standard deviations on the semantics battery were remarkably similar. The difference between the two global scores is not significant ($t(15.9) = -0.01$; $P > 0.05$).

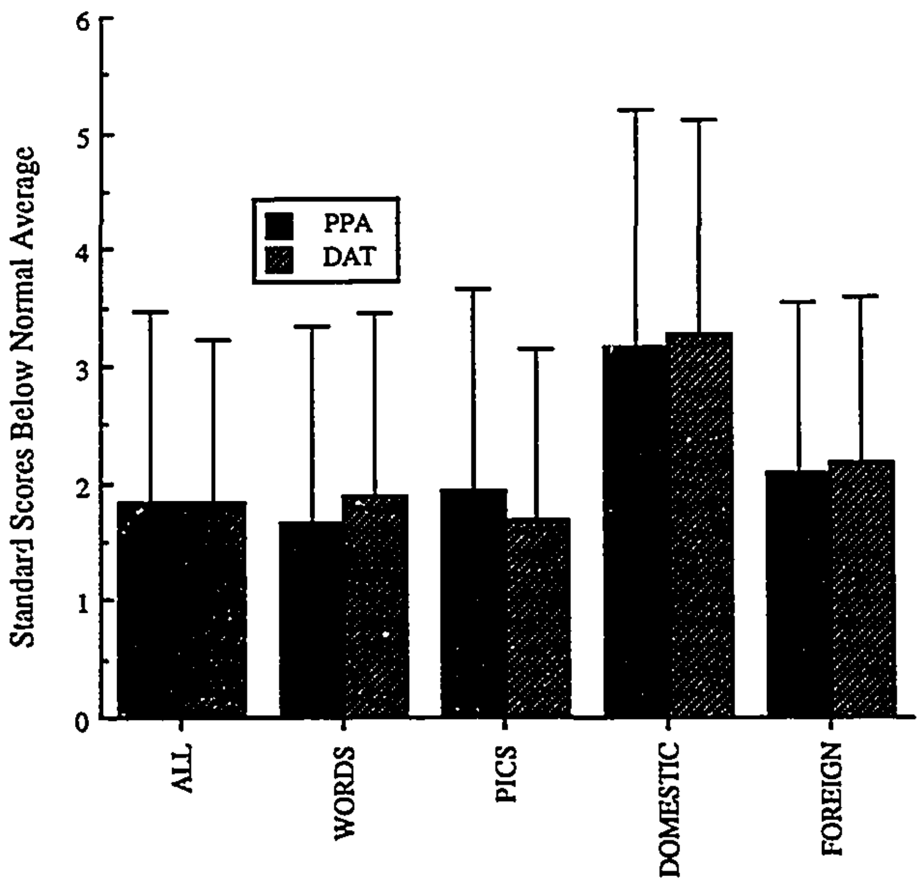


Figure 6.2
Standardized scores on the semantics battery,
by group and stimulus category.
Error bars are standard deviations.

The results from the picture naming tests are graphed in Figure 6.1. Although, as the hypothesis predicts, PPA patients did have more difficulty than DAT patients in naming pictures in both modalities, the differences were not statistically significant (oral naming: $t(16.2) = 1.59$; $P > 0.05$; written naming: $t(14.5) = 1.63$; $P > 0.05$) and, contrary to the hypothesis, the difference was larger in the written than the oral modality.

Although hypothesis 5 is not supported by the data as stated, there is one other prediction that might have been made based on the same observation: that DAT patients were more likely to have damage to pre-frontal areas, and thus more likely to have deficits in semantic search, than PPA patients. One could expect to see differences in the number of semantic errors made in the naming tests, with DAT patients making more semantic errors than PPA patients. This hypothesis accords more closely with the justification presented in Chapter 1 for hypothesis 5 than does the hypothesis as stated. The relevant raw data for assessing this new (but, alas, post-hoc) hypothesis are presented graphically in the left hand side of Figure 6.3.

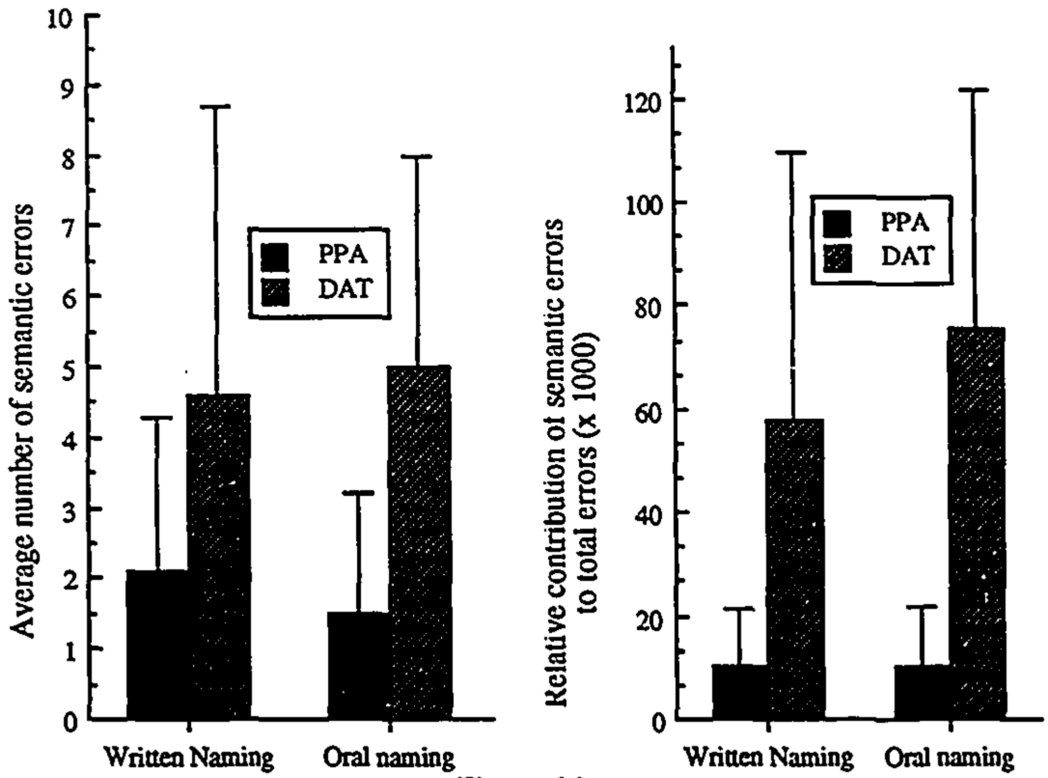


Figure 6.3
Semantic errors in naming
Left: Absolute errors
Right: Relative contribution to total errors

The difference between these two groups was analyzed in the same manner as the orthographic error data above, by calculating the contribution made to the total errors within a group by each individual's semantic naming errors, and then calculating if these proportions differed significantly between the two diagnostic groups. The averages and standard deviations of the proportionately-weighted data are presented graphically on the right hand side of Figure 6.3, which thus may serve as an illustration of why error data should be considered in this manner: note how much the group differences change when the data is represented proportionally. The difference was significant in both cases (written naming: $t(7.4) = -2.6$, $p < 0.05$; oral naming: $t(7.6) = -3.9$, $p < 0.01$).

This result emphasizes the nature of the semantic deficit in DAT, which is not a deficit preventing semantic access so much as a deficit that makes semantic access less accurate. It is thus more likely to result in a mismatch within 'semantic space' than to a failure to retrieve anything at all. Although hypothesis 5 must be rejected as stated, the significant difference in the number of semantic errors does suggest that the pre-frontal involvement of DAT may yet lead to a useful measure for functionally discriminating between the two groups.

The finding of a significant difference in the relative contribution of semantic errors to total naming errors raises the interesting question of whether or not the cortical damage associated with PPA causes a semantic deficit which is more likely to prevent access to semantics altogether (a deficit in semantic representation) than to result in a mismatch (a deficit in access of intact semantic representations), as it does in the DAT patients. This possibility may be assessed by analyzing the number of omission errors in naming in the same manner as we have analyzed the number of semantic errors. The difference in the relative contribution of errors of omission is not significant in either modality (written naming: $t(16.6) = 0.55$, $p > 0.05$; oral naming: $t(17.0) = 0.53$, $p > 0.05$).

THE MAIN HYPOTHESIS

The hypotheses presented in Chapter 1 have not fared well under examination. Hypotheses 2 and 4 have been rejected outright. Good reasons have been presented to be suspicious of the equivocal data supporting Hypothesis 3. Supporting evidence has been found for only a post-hoc version of Hypothesis 5. Because of the problematic nature of the only hypotheses which could be formulated *a priori* (as outlined in Chapter 1), these findings have more limited usefulness than the rejection of hypotheses may ideally have in science (see Popper, 1959). Nonetheless, the hypotheses have served their desired role as orienting devices for an initial foray into a complex set of data.

As a final step I must consider the status of the main hypotheses, which was stated in chapter 1 in a weak and strong version, as follows:

> Strong version: There is a systematic difference in terms of primitive psycholinguistic operations between patients with PPA and patients with Alzheimer's Disease. This difference can be stated in terms which will allow one to decide with certainty which disease a person has given that patient's performance on tests which focus on the relevant primitives.
>
> Weak Version: There are probabilistic differences in terms of primitive psycholinguistic operations between patients with PPA and patients with Alzheimer's Disease. These differences can be stated in terms which will allow one to decide with an empirically-grounded degree of probability which disease a person has given that patient's performance on tests which focus on the relevant primitives.

It is clear that no support for this hypothesis may be founded in the previous hypotheses. However, there are two other avenues worthy of exploration. The first is to review quickly the differences in the syndromes documented in Chapters 4 and 5. The second is to consider ranking the differences in performance on all tests of the PAL battery.

COMPARISON OF IDENTIFIED SYMPTOM CLUSTERS

In Chapters 4 and 5, I presented a description of the PPA and DAT patients, respectively, which focused on four main symptom clusters: the auditory input deficits, the naming deficits, the reading deficits, and deficits in affixed word production and comprehension. In this section the findings for the two groups in each of these symptoms clusters will be compared.

i.) Auditory Comprehension Deficits

Six of the eleven PPA patients had specific difficulty with parsing auditory input. Only one of the eight DAT patients showed evidence of a similar deficit. This difference in proportions is not significant (Fisher's Exact Probability: 0.19). There was a significant group difference, in favour of the DAT group, on one of the tests which is most diagnostic of an auditory input deficit: phoneme discrimination ($t(12.8) = -2.18$, $p < 0.05$). No significant differences were seen for repetition of either words ($t(17.0) = -0.88$, $p > 0.05$) or nonwords ($t(17.0) = -0.03$, $p > 0.05$). The evidence for a significant difference in auditory input processing is thus quite weak, but suggestive.

As noted in Chapter 4, there is very little known about localization of the relevant functions which can place this possible difference in acoustic processing between the groups into a framework that might allow us to make more concrete hypotheses about it. For now it must simply be noted as a possible topic for future study.

ii.) Naming Deficits

Eight of the eleven PPA patients (73%) had naming deficits. Two of these were clearly anomic, and the remainder were agnosic. Six of the eight DAT patients (75%) had naming deficits, split evenly between those considered anomic and those considered agnosic. This close accordance (Fisher's Exact Probability = 0.65) does not suggest that there are obvious systematic differences in naming between the two groups. Further evidence against such differences, as shown in the previous section, is that there are no significant group differences in any of three most directly relevant tests: oral naming, written naming, and the semantics battery.

However, we have also shown at least one difference which does distinguish significantly between the two groups: the proportion of semantic errors to total errors made in naming. DAT patients make a larger proportion of semantic errors in both modalities. As noted previously, this is consistent with the idea that damage to pre-frontal cortex may specifically impair, but not destroy, the process of mapping into semantic space.

iii.) Reading Deficits

Six PPA patients had clear reading deficits, and an additional four had more ambiguous deficits. Only one of the DAT patients had any reading deficit, and that deficit was mild. Although the difference in proportions is not quite significant (Fisher's Exact Probability = 0.06) with this small number of patients, the trend is strong enough to be noteworthy, especially when we consider that the PPA patients were much higher functioning generally than the DAT patients. Since some evidence of reading deficits was seen in almost none of the DAT patients and in almost all of the PPA patients, it seems clear that a larger group of patients would be likely to show a clear and statistically significant difference in reading ability.

The rates of reading disability in both groups accords closely with the rates reported by the reviews in Chapters 2 and 6.

iv.) Deficits in Affixed Word Processing

In Chapter 4 it was noted that every one of the PPA patients had difficulty with comprehension affixed words, and that the evidence suggested that this difficulty was not simply secondary to a more general difficulty with word comprehension and production. In contrast, not one of the DAT patients showed an unambiguous problem with affixed word comprehension. This observation of differences at the individual level is confirmed by statistical differences at the group level, which showed that the PPA performed significantly worse on all of the six relevant tests (auditory lexical decision of affixed words: $t(10.9) = 2.63$, $P < 0.05$; written lexical decision of derived words: $t(13.7) = 2.65$, $p < 0.05$; auditory word-picture matching of derived words: $t(11.5) = 3.03$, $p < 0.05$; written word-picture matching of derived words: $t(15.2) = 2.13$, $p = 0.05$; written synonym judgment of derived words: $t(11.5) = 3.15$, $p < 0.01$).

No clear analogous dichotomy was seen at the individual level among the affixed word production tests. Six of the PPA patients had a production problem which could be specifically tied to word affixation, but an argument could be made that at least three (and perhaps as many as five) DAT patients had the same problem (Fisher's Exact Probability = 0.71). However, at the group level there were significant differences in favor of the DAT group on both of the relevant tests (oral production of affixed words: $t(10.8) = 3.02$, $p < 0.05$; written production of derived words: $t(12.0) = 3.18$, $p < 0.01$).

It is difficult to draw any conclusions from this clear difference in ability to process affixed words, since neither the neurological substrates, nor the functional decomposition of this process is understood, although the analysis presented in Chapter 4 provides a potential starting point for a functional decomposition of affixed word processing. The fact that there are PPA patients who show dissociations in both directions among affixed and non-affixed word comprehension lends strong support to the view that the processing of affixed words is functionally dissociable from other word comprehension processes, perhaps especially in patients with specific deficits in recognizing nouns. This in turn makes it feasible that the process may indeed, as the evidence suggests, be differentially affected by PPA and DAT. Note too that there are several PPA patients who are very poor at morphological production and comprehension, but who nevertheless continue to function at a high level in their daily lives, and, conversely, several DAT patients who are clearly demented but able to score well on such tests. Morphological processing appears to be modularized from general cognitive function. The nature of word morphology processing in PPA is worthy of further study. More extensive studies seem likely to provide insight into the role that morphological markers play in the language system.

RANKED COMPARISONS OF SUBTEST SCORES

In order to further understand how the disease processes of DAT and PPA affect language functioning, it is useful to consider which language functions were most affected by the two disease processes by ranking tests according to 'difficulty' within each group, using the number of patients who scored significantly low as a measure of its difficulty within that group. I have ranked not only the global scores from every test, but also the scores from levels within each test. This choice simply reflects the fact that the definition of what constitutes a single test is a definition based largely on practical convenience rather than theoretical considerations. For example, we are as interested in a patient's reading of low frequency words as high frequency words, even though we asked patients to read both as part of 'a single test'. There are 129 such sets of items in the PAL battery.

Test	Items	%	> -2 SD	N
A. Comprehension- Abstract Words	ALL	100	0	8
A. Lexical Decision- Affixed Words	DER.	100	0	11
W. Production- Affixed Words	ALL	100	0	11
"	L-1	100	0	11
"	L-2	100	0	11
"	DER.	100	0	11
"	INF.	100	0	11
A. Lexical Decision- Affixed Words	ALL	91	1	11
A. W-P Mapping	HF	91	1	11
A. W-P Matching- Affixed Words	MATCH	90	1	10
Reading	W	91	1	11
"	LONG	91	1	11
"	IRR.	91	1	11
W. Comprehension- Abstract Words	ALL	91	1	11
W. Lexical Decision	NW	91	1	11
W. Picture Naming	ALL	90	1	10
"	SHORT	100	1	10
"	LONG	100	1	10
"	LF	100	1	10
"	HF	90	1	10
"	INORG.	90	1	10
"	ORG.	90	1	10
"	ANIMAL	90	1	10
"	OBJECT	90	1	10
"	FV	90	1	10
W. Synonym Judgment- Affixed Words	ALL	91	1	11
A. Lexical Decision- Affixed Words	NW	82	2	11
A. Sentence Comprehension	ALL	71	2	7
"	REV.	71	2	7
A. W-P Matching	SHORT	82	2	11
O. Picture Naming	INORG.	82	2	11
Reading	ALL	82	2	11
W. Sentence Comprehension	ALL	82	2	11

Table 6.4
Item sets failed by all, all but one, and all but two PPA patients

Almost all of the PPA patients failed 33 of the item sets, as shown in Table 6.4. Scores on tests of both written production of affixed words and auditory comprehension of abstract words were globally low for every PPA patient who completed them. In addition, every patient who was tested scored low on auditory lexical decision of derived (but not inflected) words and in writing low (but not high) frequency names. All but one patient scored globally low at auditory lexical decision of affixed words, auditory word-picture matching of affixed words, written comprehension of abstract words, written picture naming, and written synonym judgment of affixed words. Table 6.1 indicates that PPA patients had particular difficulty with reading (especially of long and irregular words), with making lexical decisions about nonwords in both modalities, with sentence comprehension in both modalities, and with oral picture naming.

Table 6.1 also points out a particular difficulty with auditory word-picture matching of short and high frequency words only. However, the reverse frequency effect is clearly an artifact, flagging a deficit in the tests rather than in the PPA patients. Since the auditory word-picture matching contains only 8 high frequency items (compared to 56 low frequency items) and normals have a very high average score at matching high frequency names to their pictures, a patient will score significantly low on that item set if he or she makes only a single error. By contrast, the length effect is not an artifact: PPA patients did indeed score significantly low far more often on matching pictures to short names than to long names. They do so because normals do quite poorly on long names as well, but make almost no errors on short words, so the short words are 'less forgiving' of errors than the long words.

Test	NAME	%	> -2 SDS	N
W. Picture Naming	LONG	100	0	8
W. Lexical Decision	W	87.5	1	8
O. Production- Affixed Words	DER.	87.5	1	8
"	ALL	87.5	1	8
"	INORG.	87.5	1	8
"	ORG.	87.5	1	8
"	OBJECT	87.5	1	8
"	FV	87.5	1	8
"	SHORT	87.5	1	8
"	LF	87.5	1	8
A. W-P mapping	SHORT	75	2	8
W. Lexical Decision	LI	75	2	8
"	REG.	75	2	8
O. Picture Naming	FV	75	2	8
W. Picture Naming	ANIMAL	75	2	8

Table 6.5
Item sets failed by all, all but one, and all but two DAT Patients

Analogous results from the DAT patients are presented in Table 6.5. Almost all of the patients scored significantly low on a total of 15 item sets.

Writing long names was the only item set with significantly low scores found for every single patient. All but one patient scored globally low at written naming as a whole, as well as at oral production of derived words only, and written lexical decision of words only. All but two patients scored significantly low on written lexical decision of low imagery and regular words, on oral naming or fruits and vegetables, and on written naming of animals. I ignore the artifact of auditory word-picture mapping of short words, for reasons explained in the last paragraph.

The only overlap between the item sets in Table 6.3 and Table 6.4 are the items from the written naming test, with the exception of the deficit at naming high frequency items, which is seen only in the PPA list (Table 6.3). The high overlap suggests one clear conclusion: that the written naming test would in itself makes a singularly inappropriate diagnostic test for distinguishing between PPA and DAT! Ten of the 24 most serious deficits which are unique to the PPA patients are associated with processing of affixed words. Another four deficits concern differences in reading. Both of these differences between the two groups have already been discussed above. Among the remaining ten deficits which are seen in the PPA group but not the DAT group, five are particularly noteworthy, because they point to two potentially dissociable functions which seem to be almost completely impaired in the PPA patients, and almost completely unimpaired in the DAT patients: sentence comprehension and comprehension of abstract words. These functions are worthy of closer analysis, and are therefore examined in more detail in the next two sections.

Sentence Comprehension

All but two of the seven PPA patients who completed the auditory and written sentence comprehension tests scored more than two standard scores below normal globally on those tests. The score for auditory sentence comprehension of reversible sentences was also significantly low for all but two of the PPA patients who completed it. In contrast, only two of eight DAT patients who completed the task scored significantly low on auditory sentence comprehension, and only three of eight DAT patients scored globally low at written sentence comprehension.

The average results from the tests of sentence comprehension are graphed in Figure 6.4. The global scores for the two groups in oral sentence comprehension are significantly different ($t(12.5) = 2.19$, $p < 0.05$), but the score for written comprehension do not differ significantly ($t(11.2) = 1.0$, $p > 0.05$). None of the group differences on sentence types are significant (oral comprehension of constrained sentences: $t(7.0) = 1.0$, $p > 0.05$; oral comprehension of reversible sentences: $t(7.2) = 1.7$, $p > 0.05$; written comprehension of constrained sentences: $t(13.7) = -0.03$, $p > 0.05$).

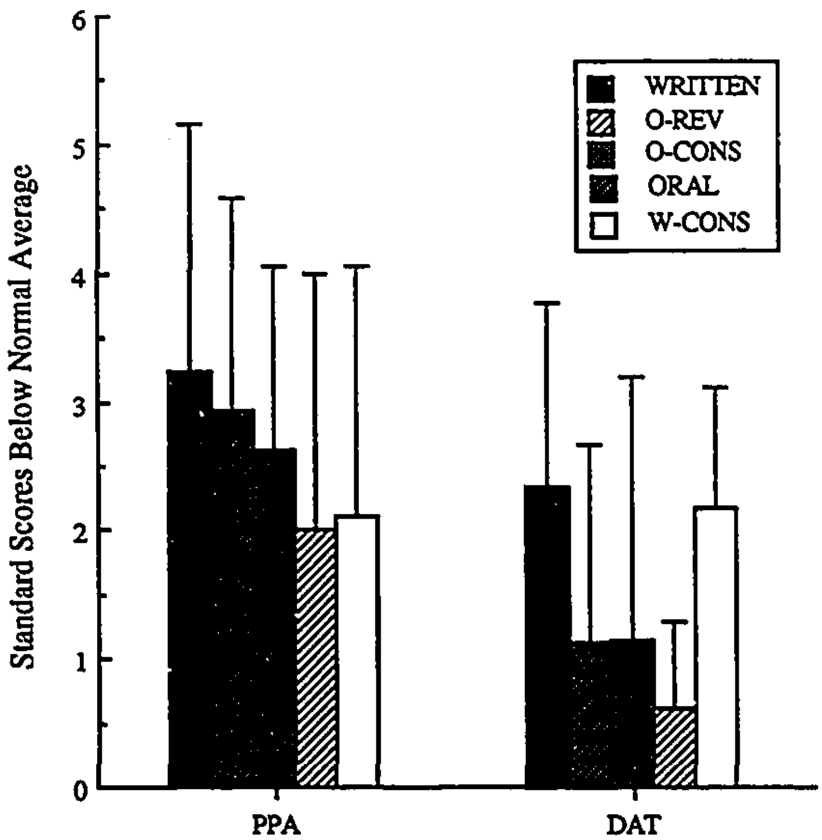


Figure 6.4
Scores on tests of sentence comprehension

In sum, the evidence for a specific impairment in PPA of sentence comprehension is equivocal. Differences at the individual level are suggestive of a particular deficit, but the small number of patients in each group who completed these tests precludes the drawing of any definite conclusions.

Abstract Words

The differences between the two diagnostic groups in abstract word comprehension are the most striking of any differences documented in this thesis. Ten of eleven PPA patients scored below norms on the written version of the abstract word comprehension test. All eight PPA patients who completed the auditory version of the test scored significantly low. In contrast, not one of the seven DAT patient who completed the written version scored low, and only two of the eight patients who completed the auditory version scored low. These differences in proportion are statistically significant in the written modality only (written abstract word comprehension: Fisher's Exact Probability = 0.03; auditory abstract word comprehension: Fisher's Exact Probability = 0.13).

The group differences, which are graphed in Figure 6.5, are significantly different (auditory abstract word comprehension: $t(12.6) = 3.07$, $p < 0.01$; written abstract word comprehension: $t(13.2) = 5.16$, $p < 0.005$).

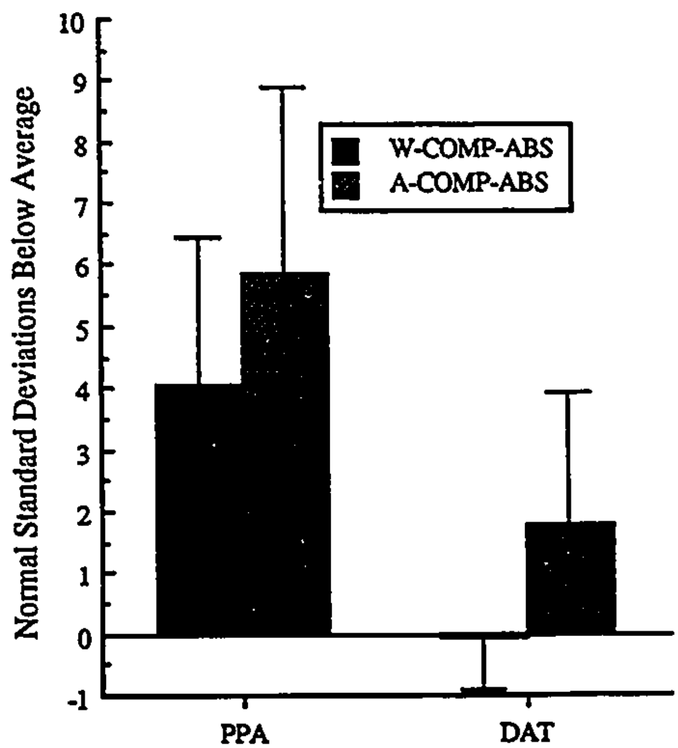


Figure 6.5

Scores on tests of Abstract Word Comprehension

The large differences at both the individual and group level on scores of abstract word comprehension, together with the consistency across modalities within both diagnostic groups, strongly suggests that PPA may affect abstract word comprehension in a way which DAT does not.

It is possible that the abstract word and affixed word comprehension deficits in PPA are related, since many of the affixed words are abstract, in the sense of being non-nouns. To test this possibility, I calculated the group Pearson product moment correlation coefficients between all pairs of the oral and auditory abstract word comprehension and synonym judgments of affixed words tests (see Table 6.5). Each test has an identical format to the test it is paired with, ensuring that any differences may be attributed to the stimuli alone. None of the correlations are significant. At least among these patients, there is no strong relation between performance at comprehending abstract words and performance on a similar test of affixed word comprehension.

	W-ABS	O-ABS	WSYN
WABS	-		
OABS	-0.47 (p = 0.30)	-	
WSYN	0.091 (p = 0.85)	0.51 (p = 0.24)	-
OSYN	0.32 (p = 0.49)	-0.01(p = 0.88)	0.57 (p = 0.19)

Table 6.5: Pearson Correlation Matrix For Tests Of Oral and Written Synonym Judgment Of Abstract & Affixed Words

CONCLUSION

In this chapter, I have compared performance of DAT and PPA on several specific test sets. My main goal in doing so has been to test my first hypothesis: that there is a systematic difference between the two diagnostic groups. Let me briefly review the findings presented in this chapter that bear on this hypothesis.

In the first section, two hypotheses were dismissed outright, as there was no evidence of relatively spared naming ability, verbal comprehension, repetition, irregular word reading, or writing in PPA patients. Only weak evidence was found to support the hypothesis that processes modulated through Broca's area (oral production of language and sentence comprehension) would be more affected in PPA than in DAT. Some evidence was presented for a post-hoc version of a fourth hypothesis which postulated a difference in the number of semantic naming errors made between the two groups. The failure to find evidence for most of the hypotheses may in part reflect the problematic nature of the hypotheses, which were all grounded of necessity on theories of functional localization, which are themselves of debatable status. A purely functional level may be a more appropriate level for theorizing, given the highly variable neuropathology in the two groups and the great difficulty of localizing linguistic subfunctions with sufficient accuracy to support hypothesizing (see Steinmetz and Seitz, 1991; Morton, 1984; and Shallice, 1988 for a list of the problems associated with theories directed at the level of localization). Unfortunately, since there was no extensive enumeration of the possible functional disruptions associated with PPA before this thesis was written, it was not possible to formulate informed a priori hypotheses at the level of functional organization.

The second section of this chapter focused on the relative performance of the patients in each group on the four symptom clusters which had been examined in some detail in Chapters 4 and 5. In that section, some evidence was presented suggesting that PPA patients are more likely to have auditory input deficits, more likely to have reading deficits, and more likely to have severe affixed word processing difficulties than DAT patients.

Finally, in the third section of this chapter, a comparison was made of performance on two other test sets, which were defined by ranking the relative performance of patients from each of the two diagnostic groups. In that section, it was shown that evidence of a specific deficit in PPA for sentence comprehension was equivocal, but that there is very strong evidence of a specific deficit in abstract word processing.

Where does all this leave the main hypothesis, which postulated the existence of definite differences between the two groups? Clearly, it is easily possible on the basis of the information presented in this chapter to design a very simple rule which will distinguish any PPA patient in this study from any DAT patient. For example, just knowing whether a patient scored significantly low on the test of written comprehension of abstract words would allow one to correctly classify all but one of the 19 patients who were considered in detail in this study. It is equally clear that it would be a trivial matter to add another single condition to the rule which allowed one to classify all patients correctly, since that condition would only need to classify the single aberrant patient. Moreover, one could, with little effort, make up any number of similar simple rules, with not more than two or three conditions each, which would have an equal ability to classify every patient in this study correctly.

The problem of taxonomy, of course, is not merely to be able to classify all the patients in a single study, but to classify all possible patients, and preferably to do so using some theoretically-motivated criteria, rather than relying on a finite set of empirical observations. Only one of the neurologically-buttressed hypothesized functional differences was found: the difference in the number of semantic errors made in naming. That difference, while presented here in a post-hoc formulation, was grounded in a large body of work showing the role played by the pre-frontal cortex in semantic access (Steinmetz & Seitz, 1991). However, many of the non-hypothesized dissociations found- between abstract and concrete words, between morphologically simple and morphologically complex words, and between sentence level versus word level comprehension- are rooted in current neurolinguistic theory, and were built in to the PAL Battery for precisely that reason (see Chapter 2).

Although I cannot yet give a confident general assessment of the status of the main hypothesis, I may at least say confidently that future work focusing on three areas- abstract word comprehension, affixed word processing in both modalities, and semantic access- is most likely to yield the information we need to evaluate this hypothesis from a theoretically-motivated point of view.

Chapter Seven: Conclusion

Difference in sameness
is the root of all delight.
Sameness in difference
is the root of all despair.

Pseudo-Heracleitus
as quoted in: Ronald De Sousa
The Rationality Of Emotion

This thesis has addressed questions in three domains of knowledge: taxonomic, explanatory, and methodological. In this chapter I will briefly review the contributions that have been made in each of these domains by the work presented in the previous chapters. I will conclude by considering how future work of a similar nature might be extended and improved.

CONTRIBUTIONS TO THE TAXONOMIC DOMAIN

The main goal of this thesis was to make a contribution towards two different taxonomic goals, one clinical and the other scientific. The clinical goal was to help specify PPA as a clinical entity, by describing the PPA and the DAT cases in as much relevant detail as was practically possible. Working towards this goal simultaneously served to work towards the more general scientific goal of creating a taxonomy of the ways the human language system can break down, independent of cause. This more general goal is one the two defining goals of the subfield of neurolinguistics. The other, of course, is the longer-term goal of explaining in theoretical terms *why* that particular taxonomy holds true- that is, explaining how the human language system actually works.

No single work can expect to 'solve' a taxonomic problem with the massive scope of these two problems. Biological taxonomists, after all, regularly look at hundreds or thousands of potential exemplars of each class before they can think of attempting to build a taxonomic system. It is to the detriment of the field that neuropsychological taxonomists cannot, for purely practical reasons, follow the example of their biologist colleagues. However, it is in part because of the limitations of neuropsychological taxonomic work, rather than despite them, that work such as this can make any contribution at all. In Chapter 3 112 PPA patients were reviewed, many of whom are described in the literature in a rather cursory manner. The number of PPA patients whose linguistic deficits have been described at a level of detail which would allow the description to be used for taxonomic purposes is a small fraction of that number. This thesis therefore contains a large percentage of the total number of the taxonomically-useful descriptions of the linguistic deficits in PPA patients that have ever been made. We should not, therefore, dismiss the utility of the work merely because it did not 'solve' the main problem that it addressed.

The ultimate utility of this work will only be apparent retrospectively, when the case descriptions which were presented here are merged with a larger group of similar descriptions. I will present a proposal of how this might be practically achieved at the end of this chapter, when I discuss future directions for this type of research. However, as I have already outlined at the end of the previous chapter, some specific contributions made to the taxonomic problem by this thesis may be identified.

Perhaps the most interesting contribution is the suggestion that PPA patients may be radically discontinuous with patients with a non-specific dementing disorder in their ability to process both affixed and abstract words. These two deficits do not appear to be related in any simple manner. Since little is known about how the language system processes either of these word types, future study of this apparent discontinuity may well shed light on the larger question of how the language system is organized. It may also provide a solution (subject, of course, to future replication of the finding) to the problem of how to identify PPAs early and with a minimal set of linguistic tests.

Although naming ability was severely affected in both diagnostic groups, a large difference was found in the proportion of semantic errors to total errors in naming. There were no concomitant differences in comprehension of semantic information. Semantic deficits are particularly difficult to assess, because (as outlined in Chapter 2) their relation to other language skills is complex. However, recent work suggests that semantic deficits may be reducible to lower-level deficits which can be measured, albeit often with some difficulty. For example, Dixon, Bub & Arguin (1994) have recently presented some evidence which proves that some semantic deficits may be more easily understood as an impairment in analyzing the abstract structure (rather than the contents) of semantic categories. Similarly, Patterson et al (1994) have recently suggested that semantic deficits and reading deficits may be more closely connected than had previously been suspected, although the data presented in Chapter 4 casts some doubt upon the details of their theory. Future progress in understanding what semantics is will allow for a much finer-grained understanding of the significance of the difference in the proportion of semantic errors in naming made by DAT and PPA patients.

Evidence has been presented in previous chapters showing a difference between PPA and DAT in auditory comprehension, reading, and sentence comprehension, although in some cases the evidence is equivocal. Closer study of these skills will hopefully shed further light on the dual problems of diagnosing and gaining a theoretical understanding of PPA.

One of the problems with studying PPA is that the term 'aphasia' is too general to be very meaningful, since it can refer to any language processing deficit (Caplan, 1992). Even the definition of what constitutes a 'language processing deficit' is becoming increasingly vague, as it becomes increasingly clear that many non-language specific deficits have implications for language. For example, low level auditory processing deficits which are not primarily linguistic may have a profound impact on language comprehension (Caplan, 1992), and, as mentioned in the last paragraph, semantic deficits which may once have seemed more cognitive than linguistic are now being shown to be closely related to wholly lexical variables (Dixon, Bub & Arguin, 1994; Patterson & Hodges, 1992). The difficulty of defining 'progressive aphasia' in a precise way is underlined by the observation, in Chapter 4, that the taxonomic clusters of the patients captured their referral histories in an almost startlingly accurate fashion. This observation raises the question of whether different diagnosticians (or their diagnostic protocol) may not be systematically sensitive to different patterns of language deterioration when diagnosing PPA. In part due to this finding, and in part due to the great heterogeneity among the PPA patients examined here, it is not clear the term 'progressive aphasia' is specific enough to be useful as a diagnostic label for the research aphasiologist (especially if s/he expects to average across groups) any more than the general label 'mental disorder' could be expected to be of practical use to the research psychiatrist. However, as outlined in Chapter 1, previous attempts to classify aphasia patients using more precise labels have met with failure for a number of reasons. Unless future taxonomic work with larger samples turns up clear regularities, we will be forced to rely on the vague term of 'primary progressive aphasia' as a general descriptive label, with the understanding that the term says almost nothing about the specific pattern of deficits of specific patient.

CONTRIBUTIONS TO THE EXPLANATORY DOMAIN

In Chapter 1 two related forms of proof for defending functional models in cognitive neuroscience were discussed: the existence proof, and hypothesis falsification. In this thesis recourse has been made to both forms of proof, in order to address some general issues having to do with the functional structure of the human language system.

The most intriguing finding which stands as an 'existence proof' was the documentation in Chapter 4 of the existence of an apparent double dissociation between affixed and non-affixed word comprehension. Two patients were found who performed consistently better at comprehension of affixed words than at comprehension of non-affixed words. At least one patient was found who showed the inverse effect. Such a double-dissociation has not been reported before. Evidence was presented which supports two possible explanations for the unusual dissociation: that it may be secondary to a difficulty processing nouns, or that it may be due to the patient's ability to glean information about an affixed word from his or her knowledge of affixes. Since it was possible to test these hypotheses on only two patients, and with a data set which was too impoverished to address the question with any measure of statistical significance, no clear conclusions could be drawn about whether either one, or both, of these explanations is correct. However, some supportive evidence was presented for both of the hypotheses, with the stronger evidence supporting the first hypothesis, that the dissociation is secondary to a specific difficulty processing nouns. An extended test of affixed word comprehension could be easily designed to test the hypothesis should another PPA patient with such a dissociation be identified.

The other form of proof- hypothesis falsification- is possible only when researchers commit themselves to making unambiguous hypotheses about the meaning of their data, a commitment which is unfortunately rare in the field of cognitive neuropsychology. Patterson et al (1992) have hypothesized that there is a functional relation between semantic memory access and the integrity of the phonological output lexicon, and that, as a consequence, a deterioration of word meaning will lead to a surface dyslexia. Data gathered for this study does not support the hypothesis.

Several patients in the study had clearly deteriorated word meaning, but were easily able to read irregular words. However, the only two patients who were clearly surface dyslexic also had deteriorated word meanings. These findings suggest that, while the directionality of the hypothesis is doubtful, there may still be a tendency for surface dyslexia and deficits in word meaning to co-occur, suggesting that the two may indeed be functionally linked.

CONTRIBUTIONS TO THE METHODOLOGICAL DOMAIN

In my opinion, the strongest contributions made by this research are the contributions to the methodological domain. It would be practically impossible to undertake such detailed descriptive work in cognitive neuropsychology without a computerized language battery, an automatic scoring system, and a method of organizing the complex data sets which needed to be analyzed. The tools developed for these tasks have already been described in detail in Chapter 2. I want to conclude by considering how those tools might be improved in order to facilitate future research of a similar nature. There are two main areas which need to be improved: organizing the increasingly complex data sets in a way which allows them to be easily mapped to theoretical models, and facilitating data collection from a larger number of patients. I will address each of these problems in turn.

ORGANIZING THE DATA SET

The only purpose of the PAL battery analysis system and the A-grams is to represent (re-present) the raw data in a way which makes it easier for experts to understand the functional dissociations in the language system that the test results document. As this thesis has illustrated, it is only at the level of functional associations and (especially) dissociations that theoretical understanding is possible. Such functional associations and dissociations can usually only be defined by looking at the results from several subtests. For example, we could not have correctly discerned the nature of the abstract word deficit described in Chapter 4 without comparing each individual's scores on the abstract word comprehension test with their scores on an analogous concrete word comprehension test. Similarly, we cannot be sure that a patient has a category specific semantic deficit unless we have evidence from several convergent sources: the word picture matching tests, the naming tests, and preferably from a semantics battery as well. My attempts to keep track of all relevant functional dissociations led me first to develop a production system which could automatically 'remember' to look for any relevant associations and dissociations between tests, and then to develop the A-grams in order to allow all identified dissociations to be graphically displayed together. In the future, I would like to merge the functionality of the production system, the A-grams, and Janus, the utility for writing rules, into a single unified system.

The simple structure of the system which I envision is diagrammed in Figure 7.1. The heart of the system would be a blank A-gram, to be used both as a template for defining rules and as a graphical output for displaying a patient's results. In order to write rules, the user could click on any one of the nodes in the diagram. Doing so would bring up a dialog box, similar to the rule-building or rule-morphing dialog box in Janus, which would allow the user to define a condition for the current rule using the data point which could be contained in that node. The user would be able to continue to define other conditions by clicking on other nodes, or to write a conclusion for the current rule, and write the rule to disk. Rules which are generated in this manner can not only be immediately incorporated into the rule set, but also, thanks to Lisp's dynamic function evaluation capability, immediately fired. This method of writing rules would continue to encourage hierarchical rule-writing in the same way that Janus did: by keeping the current rule as the default, and thus encouraging the user to change a single condition in the existent rule before writing new rules.

In order to analyze a patient's data, the user need simply choose a menu item (or click on a button on the on-screen A-gram) which would put up a directory dialog box from which the user can choose the location of the current data set, just as in the current version of the analysis system. The system will then automatically analyze that data set, fill in the A-gram on the screen, and fire all relevant rules, writing the results to a text file. The only new functionality that will be needed is the automatic completion of the A-gram, and the ability to generate rule conditions by clicking on the A-gram nodes.

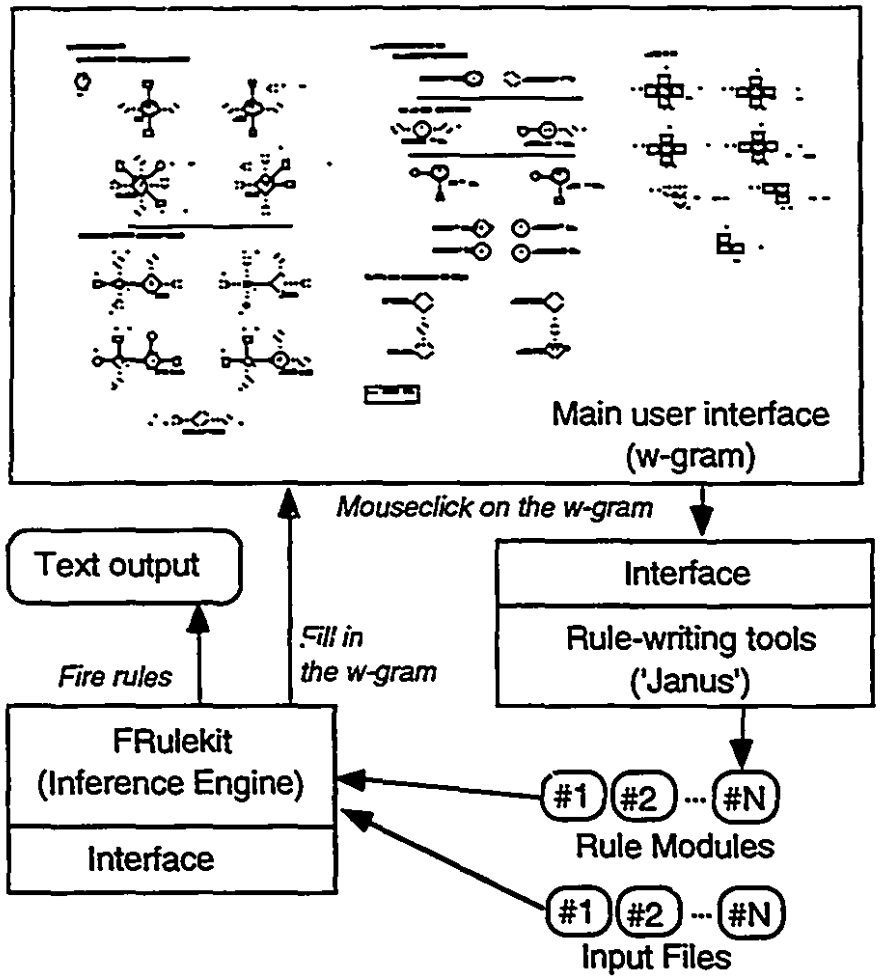


Figure 7.1
Schematic representation of a proposed system for merging
the functionality of Janus, the A-grams, and the PAL battery analysis system.
Arrows represent data flow (explained with italicized text)
Square boxes are functional modules.
Round boxes are ASCII or compiled Lisp files.
(see text for details)

Although the proposed system makes only incremental improvements to the present system (and is thus, incidentally, clearly technically possible) it will allow for much easier analysis of results, because a subject's text file and A-gram can be studied simultaneously and immediately. It will also allow for an easy way for each user to 'teach' his own system new chunks of knowledge. Because a rule can be written directly from the displayed A-gram, a user will be able, upon noticing a theoretically interesting dissociation in a patient's data for which a rule does not currently exist, to immediately write a rule to flag that dissociation in future cases, and to have that rule fire immediately, with the result written to the current result file. Note-taking thus becomes equivalent to rule-writing, with only a small increase in the effort required to make the note. Moreover, the notes which will need to be taken can be made obvious by the enhanced data visualization capabilities provided by the A-grams, which make it much easier to spot associations and dissociations in the data.

Another useful capability which may be desirable to add to the analysis system is the ability of rules to flag bibliographical references related to noted functional decompositions of the language system. It would be useful if, whenever the system identified a known syndrome, a rule would fire which would append to the patient report a bibliography of previous published work on that syndrome. The implementation of such functionality constitutes a straight-forward extension of the rule base.

Ideally it would be highly desirable to give the system access to a large data base which included the entire texts of relevant articles. The system could then ask its user s/he would like to have the entire text of any of the relevant articles appended to the patient report. This would help bridge the gap, which I perceive to be both large and widening, between clinical neuropsychology and academic neuropsychology, by giving clinicians the option to read precisely those articles which were most directly relevant to the patient they had just assessed. The addition of such functionality again presents absolutely no technical challenges, although under current copyright law there are legal problems which will make the addition of such functionality difficult to achieve.

FACILITATING THE COLLECTION OF LARGER SAMPLES

The second main problem which needs to be addressed in order to make future work on the structure and decomposition of the human language system viable is that much larger samples must be gathered and analyzed together. What the field needs is not single individuals who can gather at most a few dozen patients in the course of several years, but large groups of researchers who can collectively gather hundreds of cases per year. Such massive collaboration efforts are notoriously difficult in scientific work, in part because the coordination effort required to keep all individuals participating is overwhelming. The solution that I propose depends upon re-structuring the collaboration so that it is in each individual's self-interest to cooperate. In our society self-interest is usually measured with money. A person will typically do a task 'voluntarily' (that is, without the need for further coercion) if they will be rewarded financially to do that task. I will propose a scheme here which rewards people for their cooperation with a commodity which is rapidly becoming more valuable than money: information.

I would like to construct an explicitly copyright-free version of the PAL language battery, which would be freely distributed, along with an analysis system as described above, to any researcher or clinician who might like to use it. Distribution can now be easily effected by placing the files at a publicly accessible site on the Internet. Users of that system would be encouraged to e-mail their own rules, suggestions for improving the subtests, and (especially) their result files to a central clearing site. Rules received from users would be incorporated into the analysis system. The received suggestions would (possibly) be incorporated into the subtests. Received data would be incorporated into a database which could be either distributed with the system or used to create 'finely tuned' norms sets which would themselves be built-in to the analysis system, to allow it to score the performance of patients relative to matched peer groups. With such norms sets, the system could not only score a patient with respect to normal controls, but also with respect to (for example) age, sex, and education matched patients with a similar diagnosis. Users would be rewarded for their participation by being sent an updated version of the battery, analysis system, as well as, perhaps, the larger data base of patients. None of these items would be otherwise available, as only the 'starter' version of the system would be freely available to users who did not trade their data for updates. The data base, rule base, and analysis system would all be extremely (and increasingly) expensive to replicate, so users would have strong incentive to cooperate with the group effort.

Note that a geometrical growth in user interest may be expected under this proposal: the more people who are involved, the greater the payment for involvement will be- and, thus, the more people who will want (even if only for their own entirely selfish reasons) to become involved.

There are many other interesting methods of eliciting (and even enforcing) cooperation with such a proposal, including a number of very devious methods, which I will not detail here. I will also not address here some practical questions which would eventually need to be addressed if this proposal were ever to reach fruition, such as how one could ensure the quality of the data in the common data base, and who would get the rights to publish papers that analyze the common database. The important aspect of the proposal for the purposes here is simply the power of the two general principles: first, that each user should be paid for their cooperation with an informational tool whose worth increases in direct proportion to the total participation, and, second, that the more users there are, the more cases are collected.

It is easy and tempting to dismiss a proposal such as this one as an idealistic dream which cannot come to practical fruition. No one who has ever been involved in the development of a new set of tests can doubt that it is often nearly impossible to get group agreement on the problem of how to define of a new test. Nevertheless, I have three replies for the skeptics.

The first is that, while there will inevitably be bickering and discontent over some details, I believe that neurolinguistics has now reached the stage where nearly everyone can agree on the structure and necessity of a certain basic set of tests: roughly, the word access test of the PAL battery (phoneme discrimination, lexical decision in both modalities, reading, repetition, naming in both modalities, and word-picture matching in both modalities), plus writing to dictation. Even collaborating on collecting data for just these tests would be a huge step forward for the field.

My second reply to the skeptics is more to the point. It is an indisputable fact that it has now become so time-consuming to do a thorough analysis of a patient's language skills that it is clear that no single individual or site is going to be able to collect enough data in the span of a single working lifetime to solve the taxonomic problems which confront us. Whether it is something like my own proposal or some entirely different proposal which eventually unites the field, it is clear that the success of the taxonomic enterprise of cognitive neuropsychology is going to depend on some form of evolved cooperation.

Finally, skeptics should note that my proposal is directly inspired by a very similar proposal which depends upon the same kind of user collaboration, and which did successfully help to unite the field of child language analysis: the CHILDES project (McWhinney & Snow, 1990).

CONCLUSION

I prefaced this thesis with a quote from one of my intellectual heroes, the inventor of the essay and great refiner of the skeptic tradition in philosophy, Michel De Montaigne: "If we saw as much of the world as we do not see, we should be aware, in all probability of a perpetual multiplication and variation of forms" (Montaigne, 1575/1946, pp. 1234). This thesis has been an exercise in constraining the multiplication and variation of forms which makes up the human language system into a framework which renders that variety comprehensible. In order to make it so, I have viewed the human language system at a particular level of analysis, and with that peculiarly detailed stance which is so necessary if a viewpoint is to be dignified with the title of 'science'. It is easy to forget, while slogging through the requisite details, that all the boxes, measures, and technical minutiae which have been presented in the course of this thesis are not just dead data. Rather, they are aspects of an incredibly complex living mechanism, the very living mechanism that has allowed you to read and I to write these words. Sometimes in the course of writing this thesis I sat back and tried to imagine how all these words about damaged language systems were flowing through my own intact language system. I tried to imagine how many levels of neuronal and subneuronal systems needed to be synchronized for me to be able to decide which word would follow which, and how many millennia of evolutionary tinkering have been required to get everything to run just so. At those times I took comfort in my growing understanding that science is not just words, figures, models, and tools. It is also a privileged peek into the astounding nature of the commonplace.

The Russian novelist Fyodor Dostoyevsky once wrote (Dostoyevsky, 1880/1950, pp. 278) that "What is strange, what is marvelous, is not the idea that God really exists, the marvel is that such an idea, the idea of the necessity of God, could have entered the head of so savage and vicious a beast as man; so holy it is, so moving, so wise, and such a great honour it does man." I believe that is not just the idea of God which honours human kind. It is also the ideas of science, those same words, figures, models, and tools which have given us our peek into the astounding nature of the commonplace. It seems fitting to close this thesis with an invitation to you, my reader, to allow yourself to sit back for a moment to consider as a functioning whole the incredible language system that you are carrying around in your head. I hope that you too may come to appreciate, as I have, that it is neither mere hyperbole, nor scientific treason to say that your ability to use language is a beautiful miracle. As the biologist Melvin Konner (Konner, 1983, pp. 435) once wrote:

> "At the conclusion of all our studies we must try once again to experience the human soul as soul, and not just as a buzz of bioelectricity; the human will as will, and not just a surge of hormones; the human heart not as a fibrous, sticky pump, but as the metaphoric organ of understanding. We need not believe in them as metaphysical entities- they are as real as the flesh and blood they are made of. But we must believe in them as entities; not as analyzed fragments, but as wholes made real by our contemplation of them, by the words we use to talk of them, by the way we have transmitted them to speech. We must stand in awe of them as unassailable, even though they are dissected before our eyes."

Appendix A:
Sample Production-System Output

This appendix provides an annotated, edited example of the output generated by the PAL Battery Analysis System, which is described in Chapter 2.

Before running the system proper, all raw tests results are analyzed using a component of the production system, which produces an analysis file (listing average scores and average times for all stimuli categories, and for all pair-wise crosses of categories) for each of the tests in the PAL Battery.

The system starts (after getting the patient's name and sex) by asking for human input on four tests on which errors must be scored by linguistically-competent human beings. Scoring is facilitated by presenting the erroneous response and the target, along with a check-list of allowable error-types. A single response may be classified as exemplifying more than one error-type.

Example:

```
Please classify the errors for written-naming.

Please identify the error(s) for the following word: SLAM...SHARM
(The stimuli was: CHAIR)

   1.) Semantic Error
   2.) Phonological Error
   3.) Visual Confusion Error
   4.) Omission Error
   5.) Other Error

 >> 2

 >>
```

In this example, the patient had produced 'SLAM...SHARM' upon being asked to write the name of an item pictured on the screen- in this case, a chair. The user of the system has in this example classified this as a phonological error.

When all errors have been classified for the four tests on which such classification is necessary, the system can be left alone to score, summarize, and interpret all remaining tests. The process usually takes about 20 minutes (running on a Motorola 68030 processor at 25 MHz), and produces a report of about fifty pages.

The results from each individual test are presented in three sections:

i.) First, the raw analyzed results are echoed to the output stream. All low scores, all slow reaction times, and all small sample sizes (as occurs with some of the pair-wise crossings) are flagged with an asterisk.

ii.) A textual summary of the main results is produced, stating how the score relates to known norms for the test.

iii.) The system loads in the compiled rule file associated with the current test, and any matched rules are fired, with the relevant output written to the output stream.

Example:

```
***************************
Results for auditory word-picture mapping:

Analyzing the following file:
     11 WWP-MATCH.A 23/9/1993.

Analyzing the following file:
     04 AWP-MATCH.A 8/9/1993.

             NAME                  TIME     %     #   N

ALL                              3627.484 *  78 * 50 64
ORGANIC                          4054.375 *  75 * 24 32
INORGANIC                        3200.594 *  81 * 26 32
ANIMAL                           5391.625 *  62 * 10 16
F/V                              2717.125 *  87 * 14 16
OBJECT                           3200.594 *  81 * 26 32
SHORT                            3939.295 *  75 * 33 44
LONG                             2941.5 *  85 * 17  20
LF                               3887.268 *  76 * 43 56
HF                               1809     87 *  7   8
ORGANIC-ANIMAL                   5391.625 *  62 * 10 16
ORGANIC-F/V                      2717.125 *  87 * 14 16
INORGANIC-OBJECT                 3200.594 *  81 * 26 32
ORGANIC-SHORT                    4556 *   73 *  19  26
INORGANIC-LONG                   3396.143 *  85 * 12 14
INORGANIC-SHORT                  3048.5 *  77 * 14  18
ORGANIC-LONG                     1880.667 83 *  5   6 *
ORGANIC-LF                       4246.933 *  73 * 22 30
INORGANIC-LF                     3472.269 *  80 * 21 26
INORGANIC-HF                     2023.333 83 *  5   6 *
ORGANIC-HF                       1166     100   2   2 *
ANIMAL-SHORT                     5391.625 *  62 * 10 16
F/V-SHORT                        3219 *   90    9   10
OBJECT-LONG                      3396.143 *  85 * 12 14
OBJECT-SHORT                     3048.5 *  77 * 14  18
```

```
F/V-LONG                       1880.667 83 *  5   6 *
ANIMAL-LF                      5995.286 *  57 * 8 14
F/V-LF                         2717.125 *  87 * 14 16
OBJECT-LF                      3472.269 *  80 * 21 26
OBJECT-HF                      2023.333 83 *  5   6 *
ANIMAL-HF                      1166     100   2   2 *
SHORT-LF                       4412.694 *  72 * 26 36
LONG-LF                        2941.5 *  85 * 17  20
SHORT-HF                       1809     87 *  7   8
```

The score for auditory word-picture matching (78%) is above the cut-off point, but still below the norm.

The score for matching animal names (62%) was below the cut-off point. The score for matching object names (81%) was above that point, but still below norms. The score for matching fruit and vegetable names (87%) was within the normal range.

The score for matching short names (75%) was above the cut-off point, but still below norms. The score for matching long names (85%) was within the normal range.

The score for matching low-frequency names (76%) was above the cut-off point, but still below norms. The score for matching high-frequency names (87%) was within the normal range.

Ms. Smith is worse at matching words to pictures when animals are used as stimuli than when either fruits/vegetables or man-made artifacts are used. This pattern of response is consistent with a category-specific deficit limited to the class of animals.
****

As well as analyzing the results of individual tests, the production system has a second important task: to look for theoretically-relevant associations and dissociations between two or more tests. For the sake of human convenience, the relevant data are first re-presented in the output stream in a tabular form. The associated rule-file is then read, and any matched rules are fired.

Example:

C.) Cross-modal input results.

Task By Modality crossings:
i.) Written Versus Auditory Lexical Decision:
Result Summary (Percentage Correct) :

	Written LD	Auditory LD
Word	65	63
Nonword	100	46
Concrete	93	63
Abstract	37	63
High Freq.	75	71
Low Freq.	56	54

ii.) Written Versus Auditory Word-Picture Matching:

Result Summary (Percentage Correct) :

	Written W-P Match	Auditory W-P Match
All	76	76
High Freq.	75	87
Low Freq.	76	76

iii.) Written Versus Auditory Abstract Word Comprehension:
Result Summary (Percentage Correct) :

	Written Abstract	Auditory Abstract
Word	35	70

Ms. Smith does quite well in a test of word comprehension using concrete words. However, she fails to identify many concrete words in a lexical decision task, incorrectly rejecting these as non-words. We must conclude that the lexical decision task in this case is not providing a clear picture of Ms. Smith's ability to map auditory input onto word-forms. Ms. Smith must still be carrying out this procedure but she cannot reliably determine the status of the word, presumably due to a problem in fulfilling the requirements of a lexical decision task.

Ms. Smith is better at understanding abstract words if they are presented in spoken form than if they are presented in written form. This is a modality-specific comprehension disturbance confined to reading, a pattern that has been previously documented in the literature as being part of the syndrome known as 'deep dyslexia'.
****

In certain cases, the comparison between tests is not merely desirable, but necessary for interpreting the pattern of errors. In such cases, the relevant analysis files are consecutively read in to a single frame (without reading any associated rule files), and this frame is then written to the output stream in tabular form.

Example:

Summary Of Performance On The Semantic Battery Tests:

Results by category

		Correct	N	%
Global	-pics.	17	28	60.71
	-words	22	28	78.57
	Total	39	56	69.64
Body	-pics.	28	43	65.12
	-words	22	29	75.86
	Total	50	72	69.44
Marking	-pics.	5	11	45.45
	-words	4	11	36.36
	Total	9	22	40.91
Sound	-pics.	0	0	0.0
	-words	0	0	0.0
	Total	0	0	0.0
Enviro.	-pics.	20	27	74.07
	-words	25	27	92.59
	Total	45	54	83.33

Totals by stimuli type

	Correct	N	%
Pictures	70	109	64.22
Words	73	95	76.84

Results by animal

	Correct	N	%
bear	9	15	60.0
camel	12	15	80.0
cow	6	13	46.15
deer	10	13	76.92
dog	9	15	60.0
donkey	12	15	80.0
elephant	9	15	60.0
giraffe	10	13	76.92
horse	10	15	66.67
lion	8	11	72.73
pig	12	15	80.0
rhino	10	15	66.67
tiger	12	15	80.0

Results by domesticity (familiarity)

	Correct	N	%
North American	55	80	68.75
Non-North American	74	105	70.48

The scores for identification of names from animal markings, identification of pictures from animal markings, identification of names from animal sounds, identification of pictures from animal sounds, and identification of pictures from global characteristics were all below the cut-off point. The score for identification of pictures from animal shape was above that point, but still below norms. Scores for matching of names to animal environments, matching of pictures to animal environments, identification of names from animal shape, and identification of names from global characteristics were all within the normal range.

Appendix B: PPA Case Studies

In this appendix, the histories and PAL test results of the 11 PPA patients are presented. In order to keep the data manageable, a detailed discussion of the individual results on the tests of the PAL battery has not been included. Instead, the A-grams for each patient are presented, and only the main results are then summarized and discussed. Despite this simplification, the level of necessary descriptive detail renders these case descriptions unavoidably difficult to read. It is for this reason that they have been included in an appendix, despite the fact that these individual studies are a basic (and perhaps the most important) component of the data analysis performed in this study.

In the following descriptions, percentile scores always refer to the patient's rank within a normal, age-matched population.

INDIVIDUAL PATIENT DATA

Subject 1: AB

AB is a 72 year old, unilingually anglophone woman, with a grade 10 education. Her language problems were first noticed approximately four years before I assessed her, when she began to switch genders in speech. She now complains of having difficulty expressing herself, and findings words, and has begun carrying around a dictionary to help her. She also has difficulty understanding speech, stating that she "cannot hear" what people are saying to her, though her hearing is reported to be normal.

Due to her apparent language difficulties, standard intelligence tests were not administered. However, her pre-morbid intellectual functioning was calculated using a set of demographic variables (the Wilson Barona Formula; Barona, Reynolds, & Chastain, 1984) to fall in the average range (50th percentile). On the Mattis Dementia Rating Scale (Mattis, 1976), she scored below the normal average, although above the average scores of DAT patients. However, this test has a strong verbal loading and may have underestimated her abilities somewhat. Her score on the main non-verbal subtest (a drawing task) was above average (62nd percentile).

Memory testing did not uncover any deficits.

AB lives alone, and continues to manage her daily living activities with minimal assistance and without apparent difficulty.

Language

AB was severely impaired (over four standard deviations below the normal average) on the Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983). She had severe difficulty with repetition, and moderate word-finding and auditory comprehension deficits.

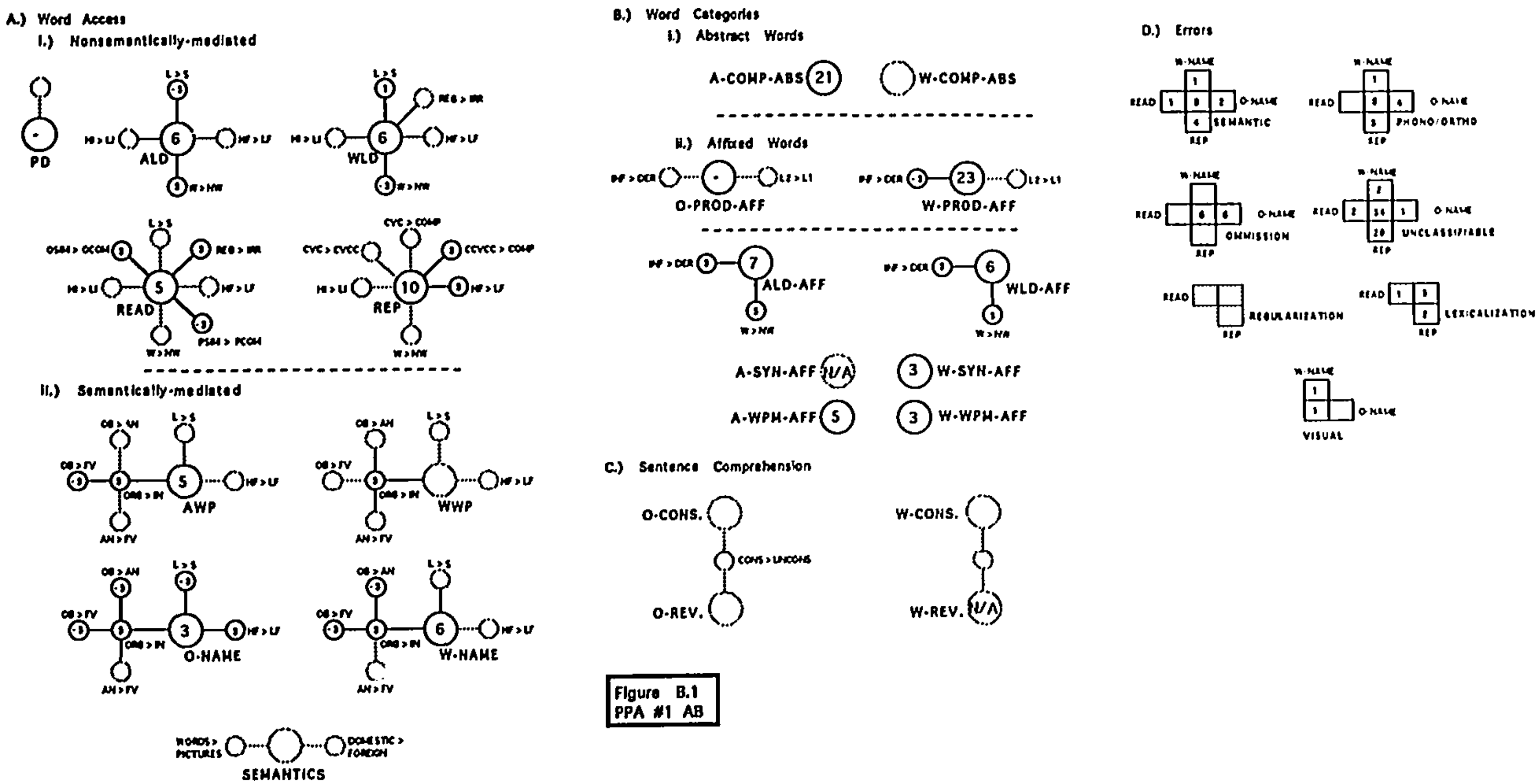


Figure B.1
PPA #1 AB

PAL Battery Results

AB's PAL results are summarized in Figure B.1. The figure may be summed up as follows:

i.) Auditory Comprehension Deficit

It is immediately obvious that AB has a great difficulty with auditory input. She was unable to complete the phoneme discrimination test, because she lexicalizes every nonword. Her performance on the repetition test (23%) was highly impaired. Although she was not significantly better at repeating words (35%) than nonwords (0%), this lack of significance was likely due to a floor effect, since she was unable to repeat even a single nonword. She was significantly better at repeating high frequency words (60%) than low frequency words (10%). Her low score on repetition contrasts markedly with her ability to read words aloud (91%), which was significantly low due to the test's sensitivity but very much better than her ability to repeat. She correctly read 94% of words, and 88% of nonwords, a remarkable feat for a woman who cannot repeat even a single nonword, and can repeat only 35% of words. AB was unable to complete the auditory test of oral abstract word comprehension, but scored within the normal range (85%) on the written version. She was also five standard deviations below norms (77%) on the auditory word-picture matching test, but within the normal range on the written version (84%). However, as we shall see in a moment, the genesis of this particular deficit is debatable.

All of this makes clear that AB is not merely unable to map subword acoustic units to pronunciation, but is unable even to correctly access the phonological lexicon from acoustic input. She has an extremely rare syndrome known as pure word deafness. Her problem probably lies in damage to the auditory input buffer, or perhaps in impairment of some sub-phonemic auditory process (see Caplan, 1992)

ii.) Anomia

AB has a mild naming deficit. She scored significantly below norms in both oral (81%) and written (84%) naming. It is tempting to assume that the difficulty is a mild anomia, since AB is within the normal range on the semantics battery. However, in this case there is some clear evidence that the problem is in fact semantically-mediated, and thus more accurately described as an agnosia: the deficit is category specific. AB is significantly worse at processing inorganic than organic stimuli on both the word picture matching tests (oral organic: 81%; oral inorganic: 72%; written organic: 100%; written inorganic: 91%) and, more clearly, on both the naming tests (oral organic: 100%; oral inorganic: 63%; written organic: 100%; written inorganic: 69%).

iii.) Deficit In Production & Comprehension Of Affixed Words

AB has a clear difficulty with producing and comprehending word morphology. Her auditory deficit made the auditory version of the affixed word production test impossible for her. She scored very low on the written version of that test (30%), making many errors which were clearly morphological in nature. For example, she wrote 'acceptine' for 'accepting', 'benefition' for 'beneficial', and 'pleasure' for 'pleasing'. She was significantly below norms on all tests of affixed word comprehension (auditory affixed word lexical decision: 67%; written affixed word lexical decision: 73%; written affixed word synonym judgment: 75%; auditory affixed word picture matching: 65%; written affixed word picture matching: 75%). The low scores on both of the lexical decision tasks reflect a difficulty in identifying nonwords only. On both tests, AB scored at chance at identifying affixed nonwords (auditory nonwords: 38%, $\chi^2(1) = 1.5$; $P > 0.05$; written nonwords: 50%, $\chi^2(1) = 0$; $P > 0.05$) but within the normal range at identifying affixed words (auditory words: 96%; written nonwords: 94%).

iv.) Deficit In Sentence Production

AB has a profound sentence production deficit. She did not produce a single correct sentence in either modality. Her errors in the oral modality are uninterpretable, since she simply failed to follow the instructions, producing sentences which were grammatical without being scorable. It is impossible to say whether this reflects a difficulty with sentence production, or a difficulty in understanding the instructions. Although it is not possible to be sure that she understood the instructions for the written tests either, her ungrammatical productions in that modality suggest a clear syntactical deficit, as well as a perseverative tendency which was not seen in any other interaction with her. Her written productions included "Woman on the door closer", "Man on the lady on her hugging for two people", and "Secretary on the disk on the book on her put desk". Her bizarre and repeated use of a single preposition suggest that it might be interesting to study her use of prepositions more closely.

Subject 2: DM

DM is a 57 year old, right handed, highly educated male. He was first seen clinically six years before we saw him, presenting with mild memory difficulties which were originally attributed to depression. A year and a half later he was shown to have evidence of 'mild dementia'. About three years after initial contact, he presented with reading difficulties. He was still extremely active, scheduling and often driving himself to his own appointments. Testing at that time showed a Dementia Rating Scale (Coblentz et al, 1973) of 120 out of 144, corresponding to mild to moderate dementia. His WAIS IQ (Wechsler, 1981a) was 86 (Verbal: 83: Performance: 92), suggesting a decline in intellectual functioning from his previous high levels. His Wechsler memory quotient score (Wechsler, 1981b) was 73.

Language

His language skills were disrupted out of proportion to these mild deficits. He was able to name only 13 and 16 of the 60 items of the Boston Naming Test on two separate administrations, although he was able to speak fluently and with good articulation. He attained a standard score of 50 (3 SDs below norms) on the Peabody Picture Vocabulary Test (Dunn & Dunn, 1982). He was surface dyslexic, impaired at reading irregular words only. Sentence comprehension was perfect. His semantic access was impaired. He was diagnosed as having semantic dementia. DM was also found to have a very rare reverse concreteness effect: that is, he was better at comprehending abstract nouns than concrete nouns. This effect had previously been described in at least one other 'semantic aphasic' (Warrington, 1975). The evidence that he has exhibited such a phenomena (and a more detailed account of his language deficits) are presented in Breedin et al, 1994.

Brain Scans

A SPECT scan approximately four years after initial presentation showed bilaterally decreased uptake in the inferior temporal lobe, more pronounced on the left. An MRI can did not uncover any abnormalities.

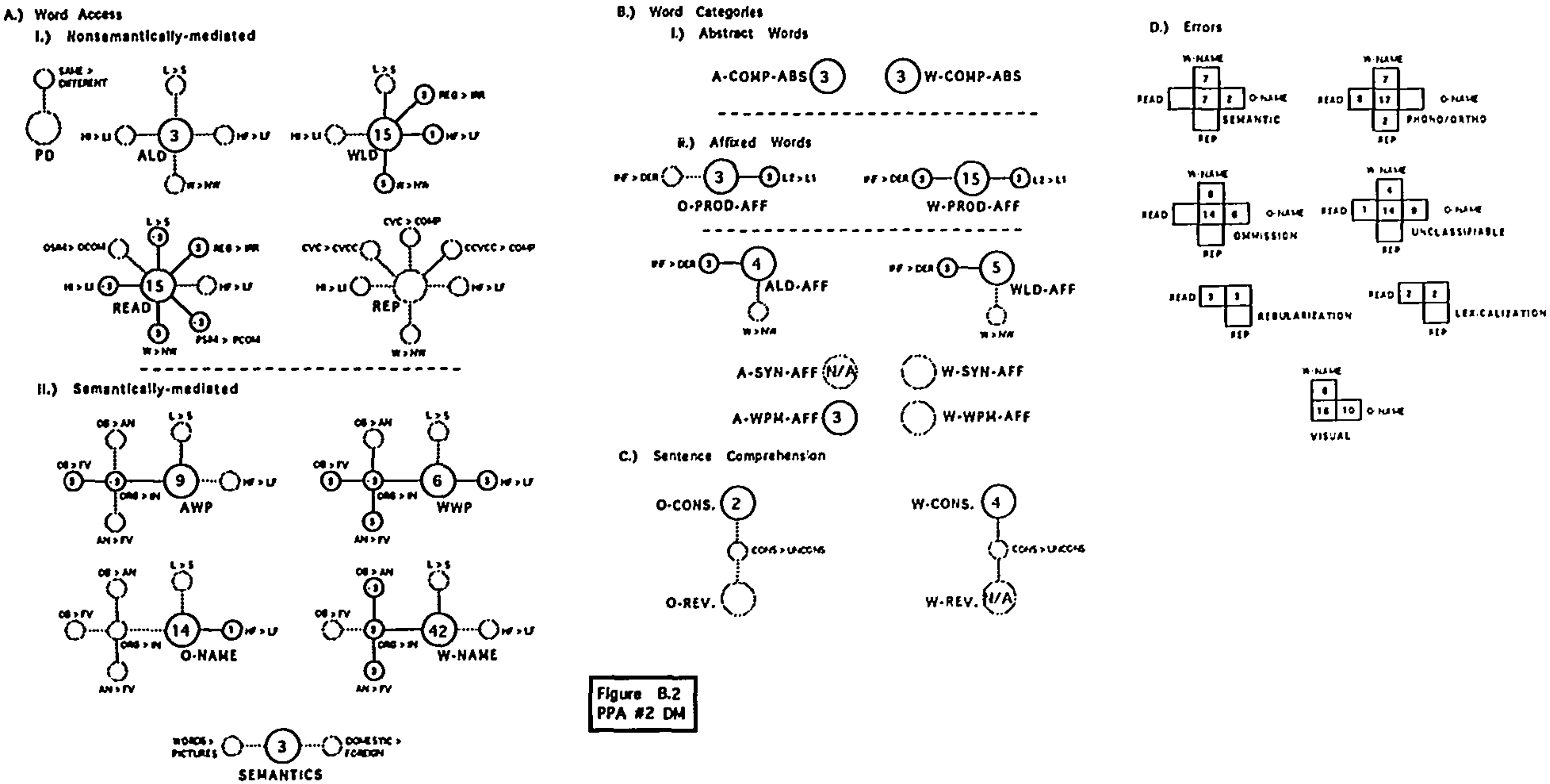


Figure B.2
PPA #2 DM

PAL Battery Results

DM's results are presented in Figure B.2. A summary of his deficits follows:

i.) Reading Deficit

DM scored significantly low on a test of reading (77%). He is worse at reading nonwords (64%) than words (89%). His word reading is significantly modulated by regularity (regular: 94%; irregular: 81%), and length (long: 81%; short: 94%).The unusual reverse imageability effect documented by Breedin et al (1994) is seen on his reading test: he scores significantly higher at reading abstract words (94%) than concrete (81%) words. (The identity of the values of the latter three dissociations is just chance.)

He also shows a reverse effect of phonological complexity: he is significantly better at reading phonologically complex nonwords (73%) than phonologically simple nonwords (50%). In light of DM's good scores on tests of phoneme discrimination and repetition, which do not support the conclusion that he is sensitive to phonemic complexity, this is a puzzling result. Without further evidence, I am inclined to regard it is artifactual.

DM's difficulty reading is also reflected in his performance on the written lexical decision test, on which he showed analogous dissociations. He is better at recognizing words (84%) than nonwords, which he recognizes at a chance level (44%; $\chi^2(1) = 0.5$; $P > 0.05$). As on the reading, he is worse at processing orthographically irregular words (79%) than regular words (93%). His performance at recognizing words on the written lexical decision task is also modulated by frequency (high frequency: 100%; low frequency: 69%).

ii.) Agnosia

DM is agnosic. He scored well below norms on all the semantically mediated tests of both comprehension (auditory word picture matching: 63%; written word-picture matching: 69%; global semantics battery score: 51%) and production (oral naming: 16%; written naming: 28%).

On the written word-picture matching test DM scored significantly higher for inorganic stimuli (78%) than organic stimuli (at chance at 53%; $\chi^2(1) = 0.1$; $P > 0.05$). Although this disparity reflects a significantly lower score in the fruits and vegetables category (at chance at 43%; $\chi^2(1) = 0.3$; $P > 0.05$) than in either the animals (63%) or objects (78%) categories, the significance of the disparity between the two organic categories is cast into doubt in this case because the score for matching animals is also not significantly different from chance ($\chi^2(1) = 1.0$; $P > 0.05$). DM was also significantly better at matching oral words to pictures in the inorganic category (72%) than in the organic category (53%; $\chi^2(1) = 0.1$; $P > 0.05$). Along with AB and CM, DM was one of only three PPA patients in this series who showed evidence of cross-modal category-specific deficit in the word-picture matching tests.

His production scores are too low to allow for any meaningful analysis of the breakdown of responses on that test. Errors on the written version of the test were about equally divided between semantic (seven errors), phonological (seven errors), visual confusion (six errors), and omission errors (eight errors), with four other unclassifiable errors. On the oral version of the test, most of the errors were visual confusion errors (ten errors) and unclassifiable errors (nine errors), although he also made two semantic errors and three errors of omission. No semantic errors were made in either reading or repeating words. This error pattern (particularly the number of omission, visual confusion, and semantic errors) lends further support to the claim that DM has particular difficulty accessing 'semantic space'.

iii.) Abstract word sparing

As mentioned above, DM has a very unusual relative sparing of abstract word reading and comprehension (as measured by his performance on the word picture matching tests) compared to his reading and comprehension of concrete words. Although he is significantly low on both tests of abstract word comprehension (oral: 85%; written 75%), his standardized written abstract word comprehension score was 6 standard scores above his written word-picture matching score, while his oral abstract word comprehension score was 11 standard scores above his oral word-picture matching score. Only one other patient (MW) showed a dissociation in the same direction, although the differences in the written modality were much less pronounced in his case than in DM's case.

iv.) Deficit in Affixed Word Production (Equivocal)

There is weak evidence of deficit in affixed word production. DM had difficulty with production of affixed words in both modalities, worse in the written (53%) than the oral (77%) modality. He was significantly better at producing level 2 words (written: 65%; oral: 85%) than he was at producing level 1 words (written: 30%; oral: 60%) in both modalities. He was also significantly worse at producing derived (40%) than inflected (80%) words in the written modality only. However, most of his errors on the oral modality were omission errors and almost all of his errors in the written modality were orthographic errors rather than true affixation errors. This error pattern makes it difficult to conclude that there is a true deficit of morphological production.

v.) Deficit In Sentence Production

DM is impaired on comprehension of semantically-constrained sentences in both the written (65%) and the oral (80%) modalities, suggesting a difficulty in processing word roles. This may simply be a function of his difficulty understanding words, as outlined in section i.) above.

DM scored 92% on the oral sentence production test, and 88% on the written sentence production test. Most of his errors were in producing sentences with relative clauses, which are very difficult to test. This performance is not suggestive of a sentence production deficit.

Phoneme discrimination and repetition are relatively spared compared to almost every other linguistic skill.

Subject 3: JD

JD is a sixty-nine year old man with a Grade 8 education. He was first seen a few months before we tested him, complaining of speech problems which had become apparent over a year earlier, and increasing in severity in recent months. When tested, he was living independently in an apartment he shared with his sister and brother.

His score on the Folstein Mini-Mental examination (Folstein et al, 1975) was 23/30, within the normal range. On the Raven Matrices test (Raven, 1965), a nonverbal measure of intelligence, he achieved a respectable score in the 60th percentile. The Digit Symbol subtest of the WAIS-R put him in the 50th percentile, although he was impaired (9th percentile) on the Block Design subtest. There was also some evidence of a memory deficit.

Language

Upon initial contact, he was noted to have major problems with verbal production and word-finding. His performance on the Boston Naming test was noted on his medical chart to be "OK". His speech was fluent and grammatical, but very difficult to understand. However, he denied having language problems, and continued to do so even after our tests had documented the extent of his difficulties, raising the question of whether he might not also have been anosognosic.

Brain Scans

EEG showed a mild, non-specific disturbance of cerebral activity. Both a SPECT and a CT scan were considered normal.

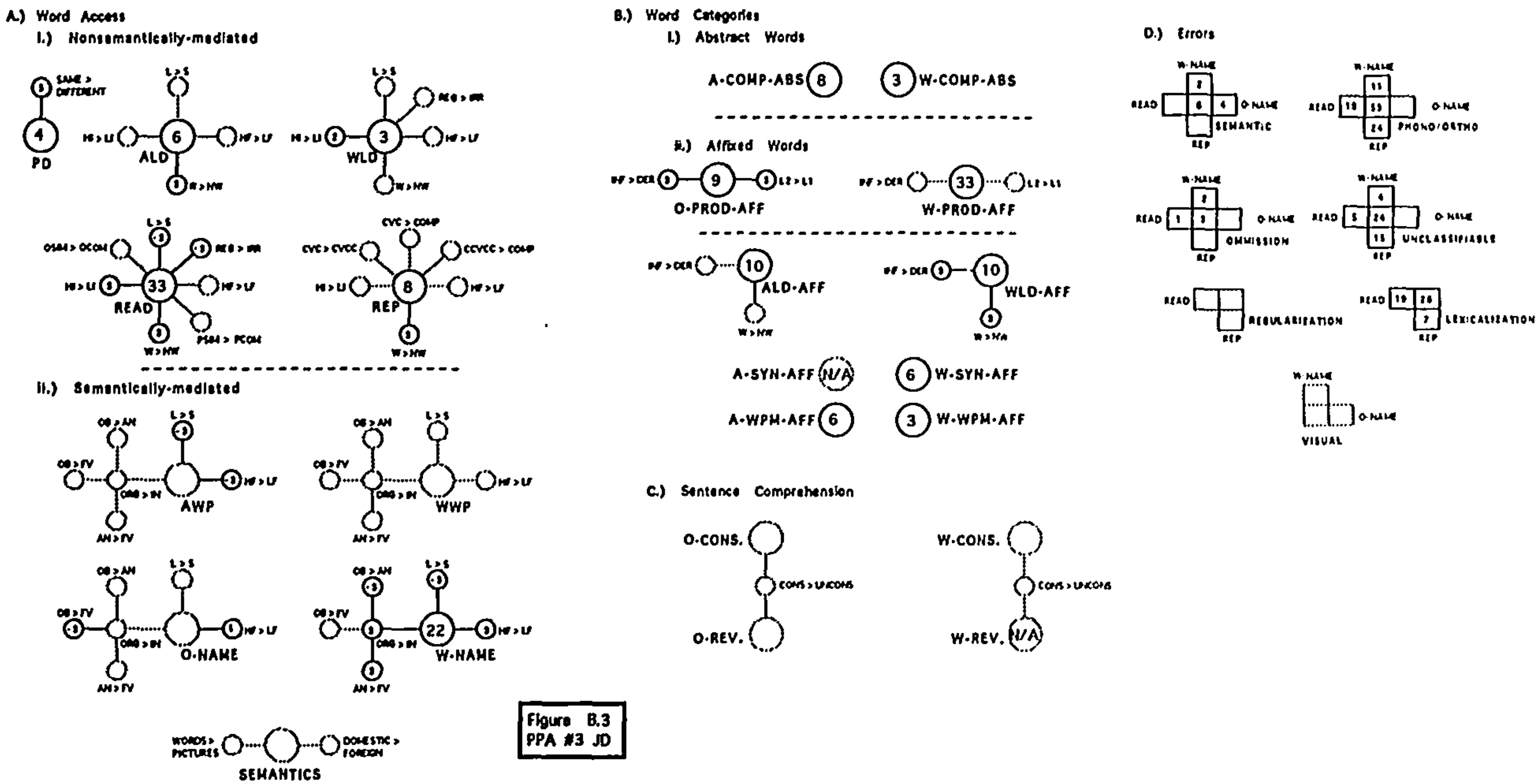


Figure B.3
PPA #3 JD

PAL Battery Results

JD's most notable deficits, as summarized in Figure B.3, are the following:

i.) Auditory Comprehension Deficit

JD has a particular problem with auditory processing, as shown by his poor performance on both phoneme discrimination (70%) and repetition (33%), his advantage in written (92%) over auditory lexical decisions (73%), and for written abstract word comprehension (75%) over auditory abstract word comprehension (at chance at 60%; $\chi^2(1) = 0.8$; $P > 0.05$).

JD's low score on the phoneme discrimination test reflects a massive response bias towards hearing the phonemes as the same. He scored 100% on repeated phonemes, but only 40% on phonemes which differed.

An analogous response bias was clear in the auditory lexical decision task. On that test, JD attained a score well within the normal range for words (97%), but was massively impaired (33%) at recognizing non-words. His near-perfect score for words and near-zero score for nonwords ensured that there were no significant dissociations in terms of stimulus characteristics. His score for recognizing written words (97%) was not significantly different than his score for recognizing written nonwords (88%).

JD was much worse at repeating nonwords (5%) than words (48%), although scores in both stimuli categories were extremely low. Most of his errors were either phonological errors (24) or unclassifiable errors (15), but he also made 7 lexicalization errors in repeating nonwords.

ii.) Reading Deficit

Unlike AB (who has a very similar set of deficits in auditory processing), JD has great difficulty reading (49%). He was able to correctly read 88% of words (significantly fewer than normals), but not a single nonword. He made many (19) lexicalization errors in reading nonwords, although the bulk of his errors were again unclassifiable (5) and phonological (18) errors. His reading of short words (94%) was significantly better than his reading of long words (81%). There was also a significant effect of imageability (abstract words: 75%; concrete words: 100%). Contrary to what might be expected from his ability to read nonwords, he read irregular words (94%) significantly better than he read regular words (81%).

This pattern of deficits (a total inability to read nonwords, with relatively preserved reading of even irregular words) is a defining characteristic of the syndrome known as phonological dyslexia (Shallice & McCarthy, 1985).

iii.) Affixed Word Production & Comprehension Deficit

Phonological dyslexics often have trouble with affixed words (Caplan, 1992). JD does indeed show evidence of a bimodal deficit in both production and comprehension of affixed words. He attained significantly low scores in every one of the affixed word tests (oral production: 47%; written production: 0%; oral lexical decision: 56%; written lexical decision: 58%; oral word-picture matching: 55%; written word-picture matching: 75%; written synonym judgment: 55%). In the oral test of affixed word production, his scores showed a significant advantage for inflected (70%) over derived (35%) words, and for level 2 words (65%) over level 1 (10%) words. Such dissociations were not possible in the written modality due to a floor effect. Significant dissociations were seen in the written affixed word lexical decision task between words (92%) and nonwords (25%), and between inflected (69%) and derived (53%) words. On the analogous auditory task, JD's score was at chance ($\chi^2(1) = 0.8.$; $P > 0.05$). Since he is utterly unable to discriminate between affixed words and nonwords, no dissociations should be expected.

JD appears to have a true morphological processing deficit. His production errors in both modalities indicate that he often has no idea of how to produce the correctly affixed form of the root word. On the written version of the test, he simply copied the root word for every item. Despite extensive prompting and checking to make sure he had understood what was required, he was unable to see anything wrong with this, and was insistent that he had written the correctly-affixed form. This is particularly interesting in light of the kinds of errors he made in written sentence production, as outlined below.

iv.) Sentence Production Deficit

JD has a bimodal difficulty with sentence production (oral: 68%; written: 24%), which may well be related to his difficulty with affixed word processing. The error pattern on both production tests suggests that JD has genuine difficulty with production of sentence structure. He orally produced sentences such as 'The ball is being thrown man to the boy', 'The grocery bag to the woman by the grocer', and 'Give the package from the little girl to the grocer'. On the written version, he seemed to have particular difficulty with conjugating verbs, writing such sentences as 'The boy is hugg the girl', 'The bag is giving by the man to the woman', and 'Boy is purry water on the plant'.

It is also noteworthy that JD's semantic access is totally spared, with his only deficit on any of the semantically-mediated word access tests (a low score in written naming) being attributable entirely to his difficulty spelling.

Subject 4: CD

CD, a 79 year old woman with a college education, was one of three patients included who do not meet the normal clinical criteria for progressive aphasia, in her case because her symptoms have not been present for two years. However, it was felt that her problems were limited clearly enough to the domain of language to justify a tentative diagnosis of progressive aphasia. She was first seen only a month before we tested her with the PAL battery. She presented with a complaint of word finding difficulties and a feeling of "knowing what I want to say, but not being able to say it". The symptoms had been present for about nine months. Although she had been educated and raised in French and had very recently started using French as her primary language again, CD had spoken English as her primary language for the last thirty years. She lived by herself, and was easily able to manage all the activities of day to day living.

Extensive psychological testing turned up only a few very mild deficits. She scored in the normal range on two dementia scales: 30/30 on the Folstein Mini-Mental test, and 163/170 on the Hierarchic Dementia Scale (Cole et al, 1983). Though her digit span was in the borderline range (12th percentile) forwards and in the low average range (26th percentile) backwards, her logical memory and immediate and delayed recognition memory were both in the normal range. Her scores on the block design subtest of the WAIS put her in the 25th percentile, and on the digit symbol in the 37th percentile. On the Yesavage Geriatric Depression Scale (Yesavage & Brink, 1983) she scored 12.6/30, in the mildly depressed range.

Language

CD's score on the Boston Naming test was 56/60, in the normal range.

Brain Scans

No brain scan data were collected.

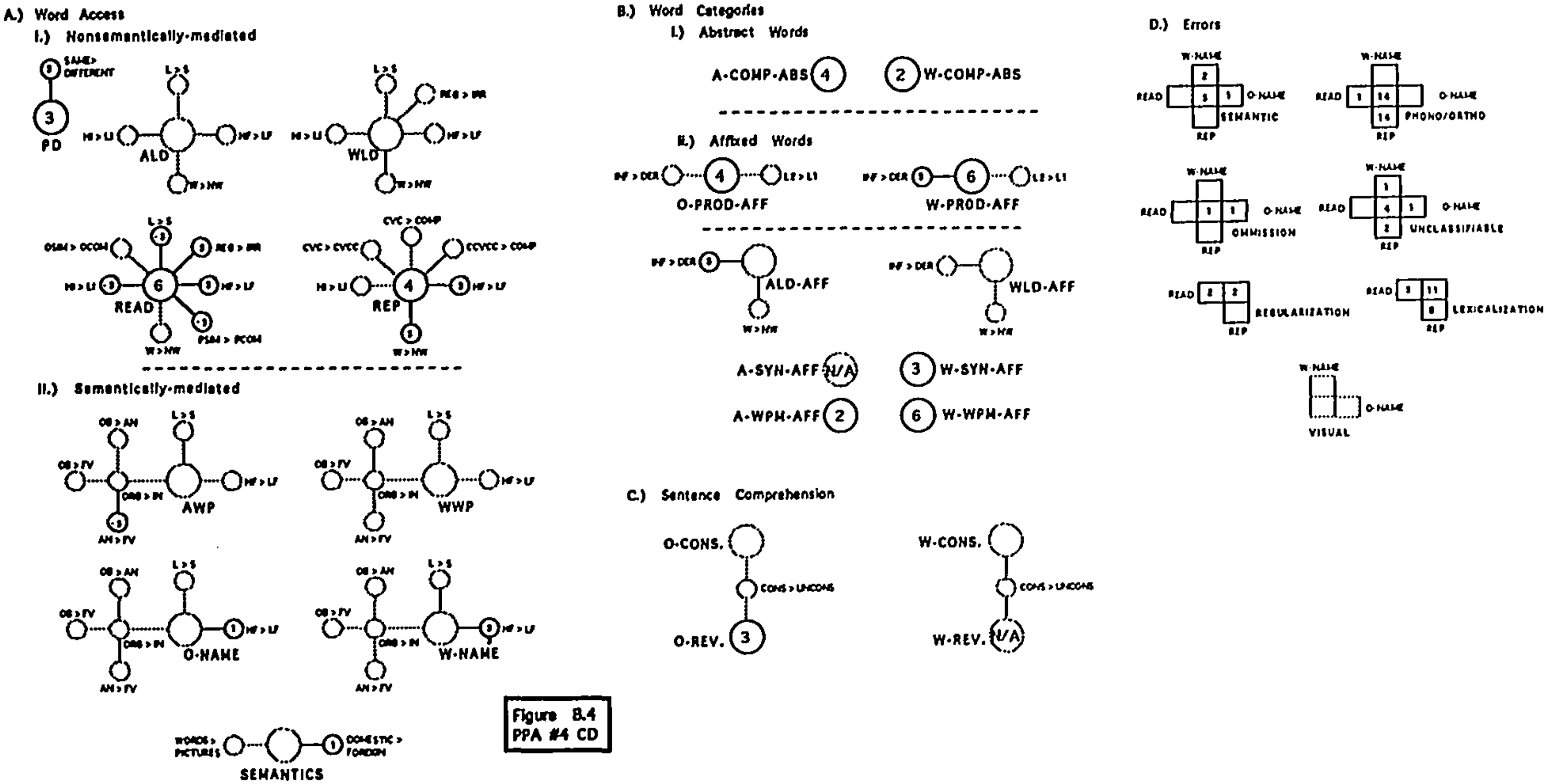


Figure 8.4
PPA #4 CD

PAL Battery Results

CD's deficits (summarized in Figure B.4) are all mild, in comparison to other patients in this study. They may be summarized as follows:

i.) Auditory Comprehension Deficit

She has signs of an auditory receptive disturbance, being significantly worse at repeating nonwords (20%) than words (85%) and tending to lexicalize repeated nonwords (8 errors out of 23). This tendency accounts for her low score (73%) on phoneme discrimination, which was due entirely to her difficulty in identifying differences (45%, compared to 100% at recognizing repeated sounds).

ii.) Reading Deficit

CD has a mild reading disturbance (89%). A mild but significant effect of lexicality was seen in the reading test (words: 91%; nonwords: 84%). There was also a mild effect of word regularity (regular: 94%; irregular: 88%). However, interpretation of this pattern of deficits must be modulated by a consideration of CD's errors. Analysis of those errors reveals that almost all of her errors were made along a dimension which is not controlled in our test: the relation of the stimulus to French words. Four out of six of her reading errors were made by pronouncing the stimuli as if it were a French word. Further testing would be necessary to be sure that this were a true reading problem.

iii.) Anomia (Equivocal)

It is noteworthy that, although her global naming scores were in the normal range, there was a significant frequency effect in both modalities, due to the fact that CD only made errors in naming low frequency items (oral low frequency: 75%; written low frequency: 76%). Since frequency is a characteristic of the stimuii names rather than the stimuli themselves, this is weak, but suggestive, evidence of an anomia.

CD's performance on all other tests of semantically-mediated word access was quite good. She achieved significantly low scores (53%) only in identifying foreign animals in the tests of the semantics battery. Her scores on the normal range on all other tests of semantically-mediated word access suggest that such access is unimpaired.

iv.) Affixed Word Production & Comprehension Deficit

She has a morphological production and comprehension deficit. Her oral production score was 73% and her written production score was 80%. With a single exception (her oral production of the nonword 'sincereness' for 'sincerity') her errors were all inappropriate, but real affixed forms of the root word.

Her comprehension deficit is more pronounced in the written modality (written synonym judgment: 75%; written word picture matching: 55%) than the auditory modality (auditory word picture matching: 82%, only just below norms).

v.) Abstract Word Production & Comprehension Deficit

CD has mild difficulty with comprehension of abstract words in both the written (80%) and oral (80%) modalities.

vi.) Sentence Comprehension Deficit

CD also has some difficulty with comprehension of oral semantically-reversible (70%) sentences.

CD's score on oral sentence production was 72%. 5 of her 7 errors were made on sentences with relative clauses, whose production is extremely difficult to test, so this probably does not reflect any deficit. On a subset of the written version of the test, CD made no errors.

Subject 5: JH

This 59 year old highly educated male is perhaps the clearest case of PPA which is presented here. As we shall see, his language skills are so globally disturbed as to make him almost untestable, while his nonverbal skills show him to be functioning at an extremely high level.

JH's wife claims to have first noticed word finding problems and stuttering (involuntary word repetition) about seven years before he was assessed for this study, although JH himself dates the beginning of his problems about four years after that, at which point his grammar and speech content began to deteriorate. By the time he was assessed for this study, JH was almost totally unable to produce speech, except for a few isolated stock phrases, such as "This one, I dunno". His speech comprehension has also deteriorated to the extent that he has great difficulty understanding normal discourse.

JH is of Hungarian extraction. He was previously fluent in English and Hungarian, and had some facility with Spanish as well. All three of these languages have been equally affected, according to his wife's report.

His most recent psychological assessment was conducted a few months after he was assessed for this study. That assessment found a Mattis Dementia Rating Scale score of 41/144. However, on the Raven's Progressive Matrices Test (Raven, 1965), he attained a score of 35/36, putting him in the top 5% of his age group on this non-verbal test of intelligence. He scored 44/54 (in the 42nd percentile) on a test of visual recognition, and well within normal limits on tests of immediate and delayed visual reproduction. He was not able to comply with the instructions on tests of verbal memory.

JH still schedules and drives himself to his own appointments. He is totally independent in his daily living activities, and, according to his wife, continues to play 'brilliant' bridge and chess. He delights in showing off his abilities, reportedly returning at one point to the second day of an assessment with a near-perfect copy of the Rey figure that he drawn from memory at home that evening. He returned to the second day of his assessment with the PAL battery with a 'cheat sheet' containing pictures of animals and their names, which he had generated at home in the evening (this is reproduced in Figure B.5).

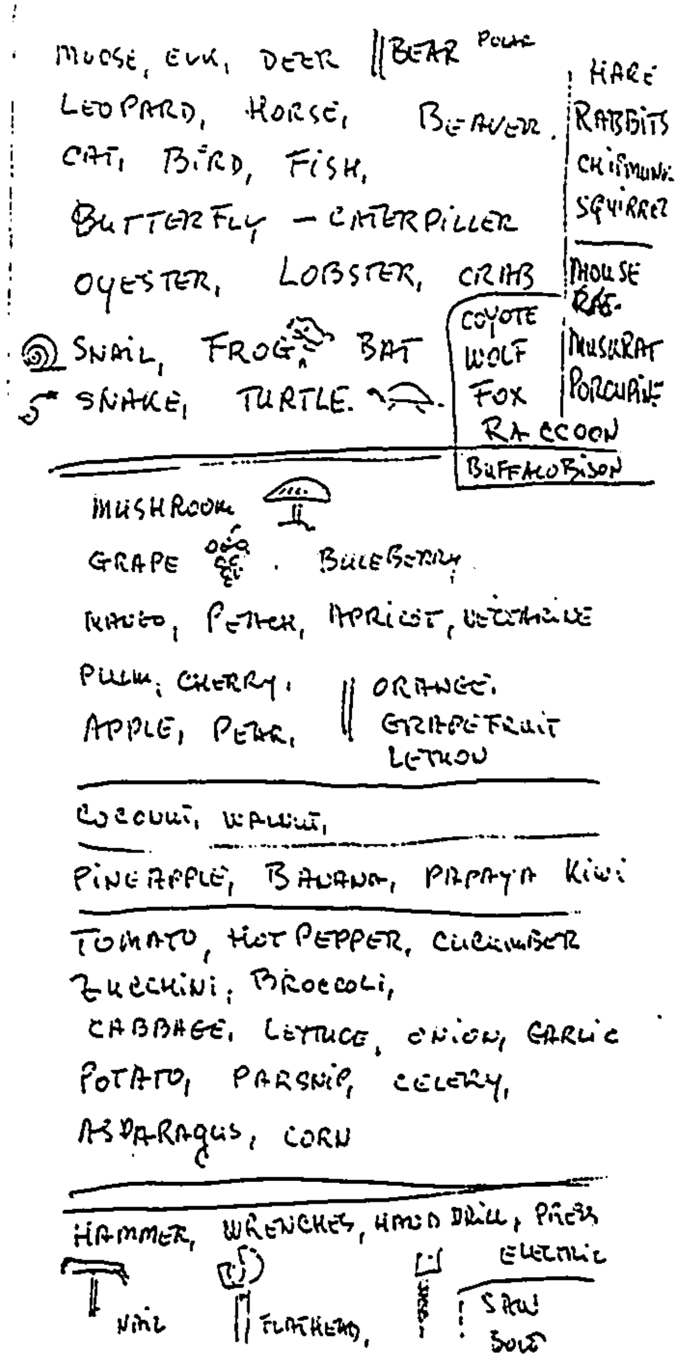


Figure B.5
Reproduction of 'Cheat Sheet' created by JH after one day of testing with the PAL

Language

The extent of JH's language deficit made it difficult to test him. He scored only 2/30 on the Boston Naming Test, and attained an Aphasia Quotient of just 10.3 on the Western Aphasia Battery (Kertesz & Poole, 1974). Moderate to severe deficits had been documented in his auditory comprehension (on the Peabody Picture Vocabulary Test, he achieved a score of 42), verbal expression, word fluency (Benton & Hamsher, 1983), and reading.

Brain Scans

An MRI revealed mild cerebral atrophy, with asymmetrical sulcal enlargement which was worst in the left perisylvian region, but which was not limited only to the left side. There was no evidence of atrophy in the frontal lobes.

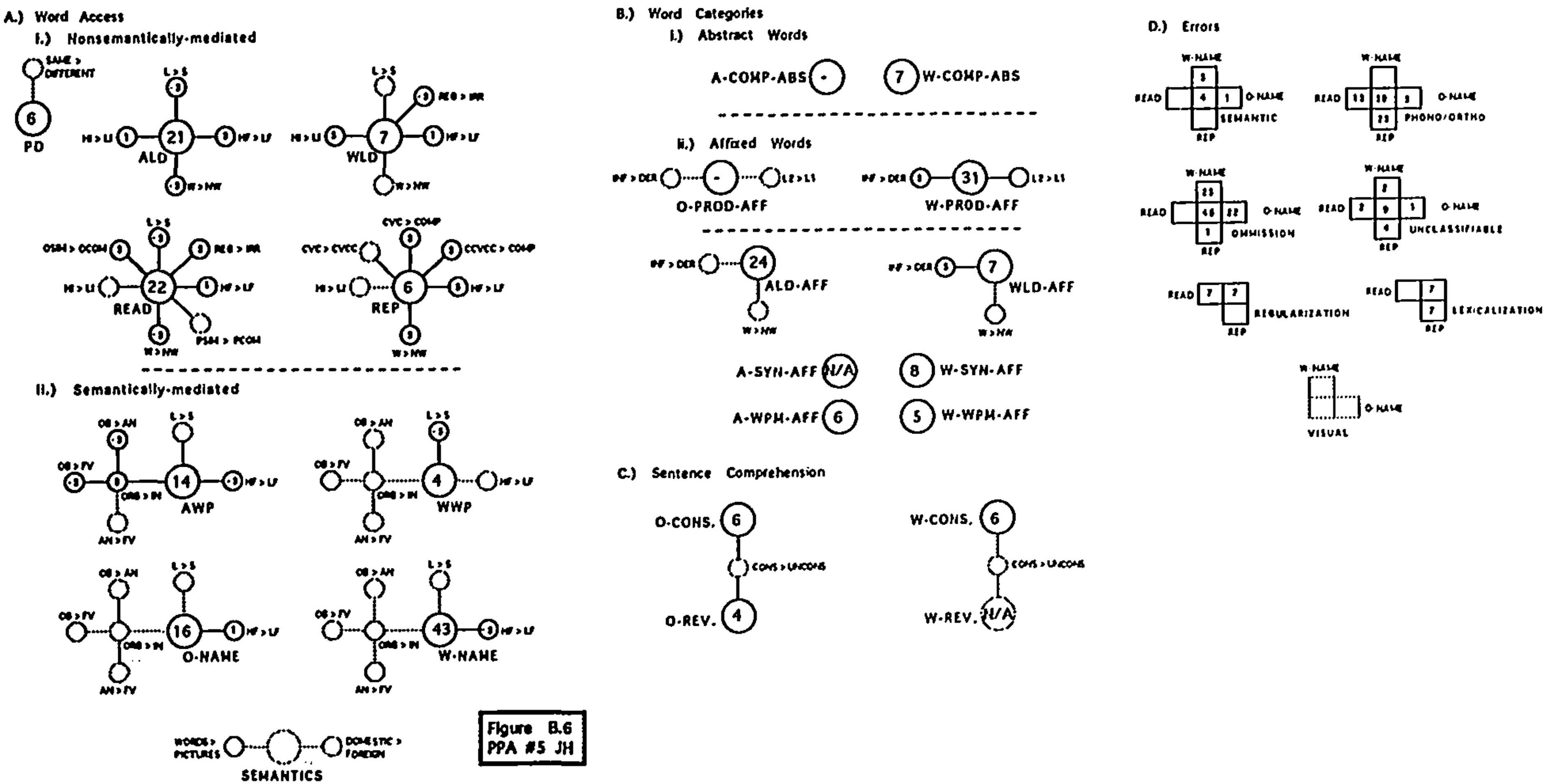


Figure B.6
PPA #5 JH

PAL Battery Results

As illustrated in Figure B.6, JH scored below norms on every test in the battery, except for the tests of the semantic battery, on which he scored low for domestic animals only. Given such widespread impairment, it may seem useless to try to summarize his deficits. However, some interpretation is made possible by the patterns of deficits. The following claims have some support:

i.) Auditory Comprehension Deficit

JH's processing of auditory input is much more impaired than processing of written input. Although he was generally very cooperative, he adamantly refused three times to complete either the auditory lexical decision task or the phoneme discrimination task, signaling after a few stimuli that this test was impossible for him. Only after the strongest insistence did he reluctantly agree to complete the two tests. His performance on the phoneme discrimination task (55%) was not significantly better than chance ($\chi^2(1) = 0.4$; $P > 0.05$), with no significant differences on judging stimuli which were the same (65%) or different (45%). His performance in auditory lexical decision (22%) was below chance. He scored 0% at identifying long (but not short) words, low (but not high) frequency words, and abstract (but not concrete) words. He was significantly better at recognizing nonwords (33%) than nonwords (17%).

Although JH was also well below norms at lexical decision in the written modality, he was able to make decisions in this modality at a rate (81%) better than chance ($\chi^2(1) = 0.4$; $P < 0.01$). He was better at making decisions on nonwords (84%) than on words (78%), the difference attaining significance due to the small normal discrepancy between these two categories. He was also significantly better at making decisions on orthographically irregular words (89%) than orthographically regular words (64%), better at making decisions on concrete (94%) than abstract (63%) words, and much better at making decisions on high frequency words (100%) than low frequency words (56%), though we do not have statistical significance measures for this latter discrepancy because we do not have data on normal performance.

JH is severely impaired at repetition (52%), especially of nonwords (15%, compared to 70% for words), and low frequency words (50%, compared to 90% for high frequency words). The bulk of his repetition errors (21 of 23) are phonological and lexicalization errors. Within the nonword category, there is a clear and significant distinction of his ability to repeat orthographically complex nonwords (13%) from his ability to repeat CVC stimuli (63%) or CCVCC stimuli (67%). The differential performances long the lines of stimuli characteristics proves that, notwithstanding the phoneme discrimination and auditory lexical decision results, some processing of auditory stimuli is occurring. JH's problem must not lie at the level of the auditory input buffer, but must rather be due to a partial, but severe, degradation at the level of the auditory input lexicon. His almost total inability to repeat nonwords implies that he is not able to use the 'sub-word level auditory-to-phonological conversion route'. The implication is further supported by his tendency towards lexicalization errors in repeating nonwords.

JH's auditory impairment is also apparent from his score at matching words to pictures in the auditory modality (44%), which was not significantly different from chance ($\chi^2(1) = 1$; $P > 0.05$). In the written modality, he was much better (75%)

ii.) Reading Deficit

JH has a reading deficit (67%). He was better at reading nonwords (72%) than words (62%), replicating the effect found in the written lexical decision task. JH is very much better at reading orthographically simple words and nonwords (93%) than orthographically complex words and nonwords (40%). He is also much better at reading regular words (75%) than irregular words (50%). This pattern of deficits is the defining characteristic of the syndrome known as surface dyslexia.

Within the words, JH was also significantly worse at reading long words (56%) than short words (69%). There was a clinically-significant frequency effect (low frequency words: 47%; high frequency words: 76%).

His reading errors are idiosyncratic, consisting largely of his insistence on pronouncing the silent 'e' at the end of many stimuli. For example, he reads 'veena' for 'vine', 'flop-a' for the nonword 'flope', and, rather remarkably in light of the word's notorious difficulty, 'tortissa' (instead of 'tortoiza', as one might have expected) for 'tortoise'. When there is not a silent 'e' at the end of the word, he never adds the syllable.

iii.) Anomia

JH has a naming deficit. As mentioned, he could not match auditory words to pictures at a rate better than chance, but was able to match some written words to pictures (75%). The only significant effect of a stimuli characteristic in was a disparity in matching pictures with short names (80%), compared to matching pictures with long names (60%), the same effect which had been seen in reading and auditory lexical decision.

JH's performance on the semantics battery is mainly within the normal range, with only a mild deficit (69%) in accessing information about domestic animals. In light of his difficulty in mapping written animal names to pictures on the binary choice written word-picture mapping test, it is interesting to see his ability to match animal characteristics correctly from animal names on the four-choice semantics tests. This suggests that his naming problem is due more to a difficulty accessing names than to a difficulty with accessing semantic information.

JH is almost completely unable to name visual stimuli in either the written (13%) or the oral (16%) modalities, despite his insistence upon having access to his 'cheat sheet' for both tests. His real ability to name is certainly somewhat lower than these scores suggest. In both modalities, he was better at naming items with high frequency names than he was at naming items with low frequency names (25% to 6% in the oral modality, and 18% to 7% in the written modality). However, this result must be considered untrustworthy, since it may simply reflect a frequency effect on his cheat sheet.

iv.) Abstract Word Comprehension Deficit

JH does not seem able to comprehend abstract words. Results from the auditory abstract word comprehension test were not obtained, as JH was unable to complete this task because of the difficulty with auditory processing documented above. On the written version of this binary choice test, he was at chance (50%; $\chi^2(1) <= 0$; $P > 0.05$).

v.) Affixed Word Production & Comprehension Deficit

JH has real deficit in comprehension and production of affixed words.

Comprehension of affixed words was very impaired. JH scored below norms on all three tests of affixed word comprehension in both modalities. He was almost completely unable to complete these tests in the auditory modality (7% on lexical decisions and 59% on the binary-choice word-picture matching test). In the written version of the affixed word lexical decision test, he achieved a score (69%) above chance ($\chi^2(1) = 6.8$; $P < 0.01$) and showed a significant advantage for recognizing inflected (94%) over derived (56%) words.

JH was not able to produce responses on the oral version of the affixed word production test. He had great difficulty (just 7% correct) with the written version of that test, producing only two correct responses, both inflections. His error pattern suggests that he has a true morphology deficit. For example, he produced the over-regularized forms 'writed' for 'wrote' and 'falled' for 'fell'. Most of his other errors were allowable affixed forms which did not make sense in the context of the cue sentence. For example, he wrote 'collides' when the answer was 'collision', 'opposed' for 'opposition' and the incorrect form 'flys' for 'flew'.

vi.) Sentence Comprehension Deficit

JH has a serious sentence comprehension deficit. He was at chance ($\chi^2(1) <= 1.6$; $P > 0.05$) on all four subsections of the two sentence comprehension tests.

He was not able to attempt the sentence production tests.

Subject 6: BH

BH is a 74 year old female with a Grade 11 education who presented with complaints of difficulty speaking. A year after initial presentation, when she was assessed using the PAL battery, she described her problem (in terms almost identical to the way CD described hers) by explaining that she sometimes 'has a feeling inside of what I want to say, but I can't say it'. She was still living independently at that time. According to the Yesavage Geriatric Depression Scale (15/30), she was mildly depressed.

In her initial assessment, her scores on the Folstein Mini-Mental status test (22/30) and the Hierarchic Dementia Scale (155/170) suggested some mild cognitive impairment. Her performance on the Wechsler story recall and forward digit span were in the low average range. She was impaired at reverse digit span and on the delayed recall of the Rey Auditory-Verbal Learning Test. On the Digit Symbol subtest of the WAIS-R she was in the low average range (16th percentile). On the Block Design subtest she was in the mildly impaired range (5th percentile).

Language

At the time of the first assessment, BH was noted to make paraphasic errors. Word fluency was impaired, but her score on the Boston Naming test (56/60) showed that naming was not impaired. She was noted at that time to make paraphasic errors in her speech.

Brain Scans

A CAT scan soon after presentation was normal. EEG suggested a left temporal irregularity. A SPECT scan found of decreased blood flow in the left perisylvian region. An MRI revealed subcortical irregularities only, with lacunae in the basal ganglia and corona radiata.

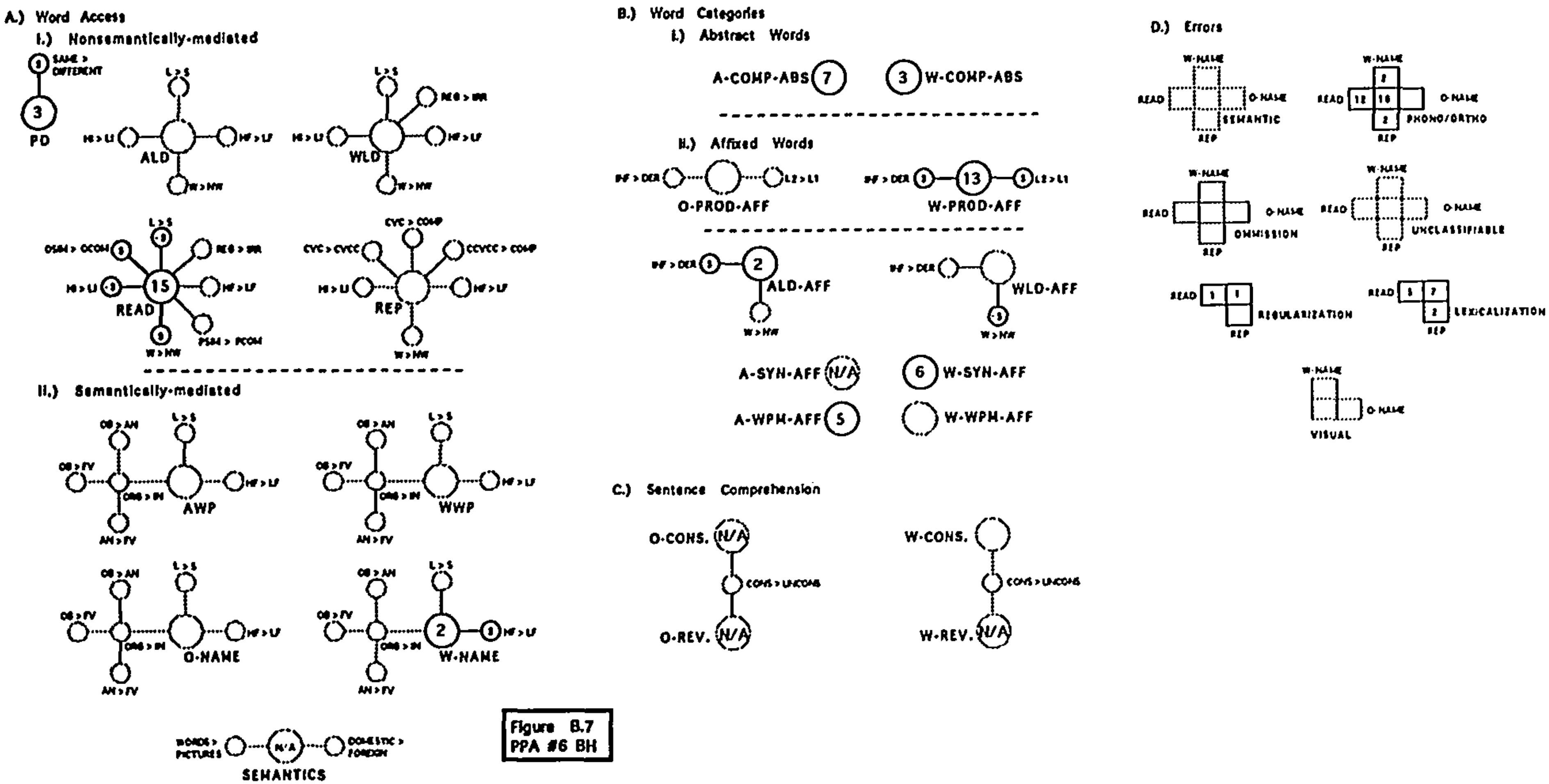


Figure B.7
PPA #6 BH

PAL Battery Results

BH's diagnosis is questionable, since she was assessed using the PAL Battery just one year after presentation, and since there is evidence that she had at that time extra-linguistic cognitive deficits, as described above. Our assessment turned up three main deficits (see Figure B.7):

i.) Reading Deficit

The main deficit is a clear phonological dyslexia. BH read only 56% of nonwords, compared to 94% of words. Her difficulty is made clearer by noting how her nonword reading errors break down: she was able to correctly read only 30% of orthographically complex nonwords, compared to 73% of orthographically simple nonwords. Clearly she is relying upon constraints provided by lexical access in order to convert from written input to phonological output.

BH was also significantly better at reading short (100%) than long (88%) words, and at reading low imagery (100%) than high imagery (88%) words. Given the relatively high scores in all four categories (compared, for example, to the dissociations documented in the last paragraph) I am disinclined to put much weight on these dissociations, which are probably just reflections of the over-sensitivity of the reading task.

ii.) Auditory Comprehension Deficit (Equivocal)

There is some evidence that BH has a modality-specific comprehension deficit. She is at chance on both auditory comprehension of abstract words (65%; $\chi^2(1) = 1.8$; $P > 0.05$) and on auditory word-picture matching of affixed words (50%; $\chi^2(1) = 0$; $P > 0.05$). She also scored significantly low on auditory lexical decision of affixed words (88%). In contrast, she scored above chance on written comprehension of abstract words (75%; $\chi^2(1) = 5$; $P < 0.05$), and within the normal range on written affixed word picture matching (85%). However, her written comprehension was not perfect: she was at chance (55%; $\chi^2(1) = 0.2$; $P > 0.05$) on written synonym judgment of affixed words.

iii.) Affixed Word Production Deficit (Written Only)

BH also has difficulty with written production of affixed words (60%), with her errors on that test and pattern of results on the other tests supporting the inference that the problem really is in finding the proper word affixation, rather than being secondary to a more general problems with orthographic output. For example, she offered as responses the nonwords 'acceptant' and 'extendent'. Her scores also suggest a real morphology problem, since they perfectly reflect the theoretical difficulty of correctly affixing a word: she is significantly better at producing the correct inflected word (80%) than derived word (50%), and significantly better at producing level 2 derived words (75%) than level 1 derived words (30%).

BH's mild written naming deficit reflects only one true naming error 'telesphone' for 'telescope'. Both her other errors were pluralization errors. This pattern of errors does not justify the inclusion of a naming deficit as being among her symptoms.

Subject 7: BL

BL, a 70 year old unilingually-English female with a college education, presented herself in 1992, complaining of a word-finding problem which had been growing worse for some months. The problem reached crisis proportions for her when she found herself unable to reliably name the card suits, a loss which forced her to give up reluctantly her passion for playing bridge. She continued to be involved in the family business, drove herself to and from appointments, and maintained an active and independent lifestyle.

She scored within the normal range on the Folstein Mini-Mental exam (27/30), the Hierarchical Dementia Scale (165/170), and on the Block Design (37th percentile) and Digit Symbol (50th percentile) subtests of the WAIS-R. However, she had some memory deficits, being totally unable to recall either story from the Wechsler Memory Scale, and scoring below average on immediate and delayed recall of the Rey Auditory-Verbal Learning test.

She was mildly depressed (8/30) on the Yesavage Geriatric depression scale.

Language

BL's scored significantly low (20/60) on the Boston Naming test. Her word fluency was in the low average range.

Brain Scans

An EEG a year prior to her language assessment revealed a mild irregularity over the left (and, to a lesser extent, the right) temporal lobe. A CAT scan was normal. An MRI scan one year later showed cortical atrophy consistent with age.

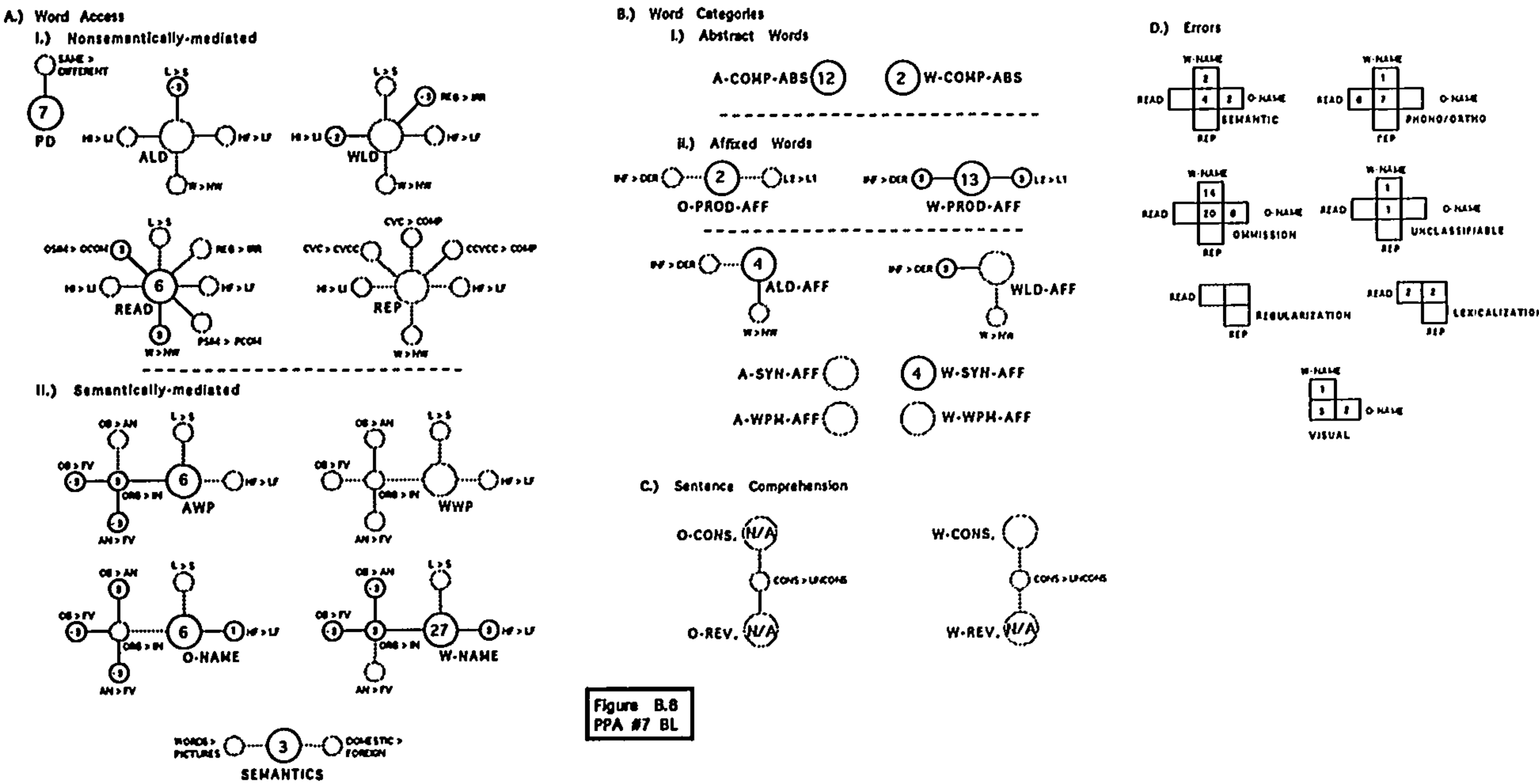


Figure B.8
PPA #7 BL

PAL Battery Results

An A-gram showing BL's results is presented in Figure B.8. The result may be summarized as follows:

i.) Reading Deficit

BL is much better at reading words (100%) than nonwords (76%), and better at reading orthographically simple nonwords (93%) than orthographically complex nonwords (50%). This pattern of deficits, indicating interference in converting from subword level orthography to phonology, is commonly know as phonological dyslexia.

ii.) Agnosia

BL is seriously anomic, scoring well below norms on both oral (66%) and written (44%) naming. Only one of nineteen errors on the written naming task is attributable to orthography, indicating that there really is greater interference in accessing names in the written modality. Her poor performance on the semantic battery suggest that her difficulty in finding names is due to a deficit in accessing semantic information.

The naming error seems to be modulated by semantic category, which lends further support to the inference that BL has a semantic deficit. In the oral modality, fruits and vegetables (100%) were named significantly better than either animals (25%) or objects (69%), a pattern which was also seen in the auditory word picture matching test. On the written naming task, BL scored significantly lower at naming inanimate objects (25%) than fruits and vegetables (63%). However, her score at naming animals (63%) was not lower than her score at naming fruits and vegetables. Although the pattern is not perfectly consistent, there is consistency in the sparing of access of semantic information about fruits and vegetables relative to access of information from the other categories.

iii.) Auditory Comprehension Deficit

There is evidence of a modality specific comprehension deficit limited to audition. BL was notably worse in the auditory modality on every comprehension task for which a bimodal comparison is possible. BL scored below the cutoff point on auditory (75%) but not written (97%) word picture matching, and on auditory (81%) but not written (90%) lexical decision of affixed words. She was at chance on the test of auditory (45%; $\chi^2(1) = 0.2$; $P > 0.05$) but not written comprehension of abstract words (80%; $\chi^2(1) = 7.2$; $P < 0.01$).

This difficulty in parsing auditory input may explain BL's chance performance (52%; $\chi^2(1) = 0.1$; $P > 0.05$) on the phoneme discrimination task. Her poor performance on this task suggests that the auditory problem must be quite low level.

iv.) Affixed Word Production Deficit

BL also has a deficit in affixing words. She was below norms in both oral (83%) and written (60%) affixed word production. Two pieces of evidence suggest that this result reflects a true morphological problem. The first is her pattern of errors: on the written test, she was significantly worse at producing derived (50%) than inflected words (80%), and significantly worse at producing level 1 inflected words (20%) than level 2 inflected words (80%). The second piece of evidence of that some of her errors are clearly at the level of word morphology rather than orthography. For example, she produced 'righten' for 'wrote' and 'sincerenes' for 'sincerity'.

Subject 8: JL

JL is a 70 year old, college-educated, right handed male with an 8 year history of progressive language loss whose condition had deteriorated into general dementia by the time we assessed him for this study. His presenting complaints were word-finding problems and expressive difficulties. In recent years he has shown soft signs of frontal lobe involvement, becoming more withdrawn, stubborn, and easily annoyed. He can follow only simple one-step commands, but is unable to calculate, and now incapable of looking after himself.

The lack of psychological test results and JL's advanced dementia at this stage render his status for inclusion in this study uncertain. However, on the basis of his reported history, we included him to see if there were patterns of spared performance which might shed light on the process of his language dissolution.

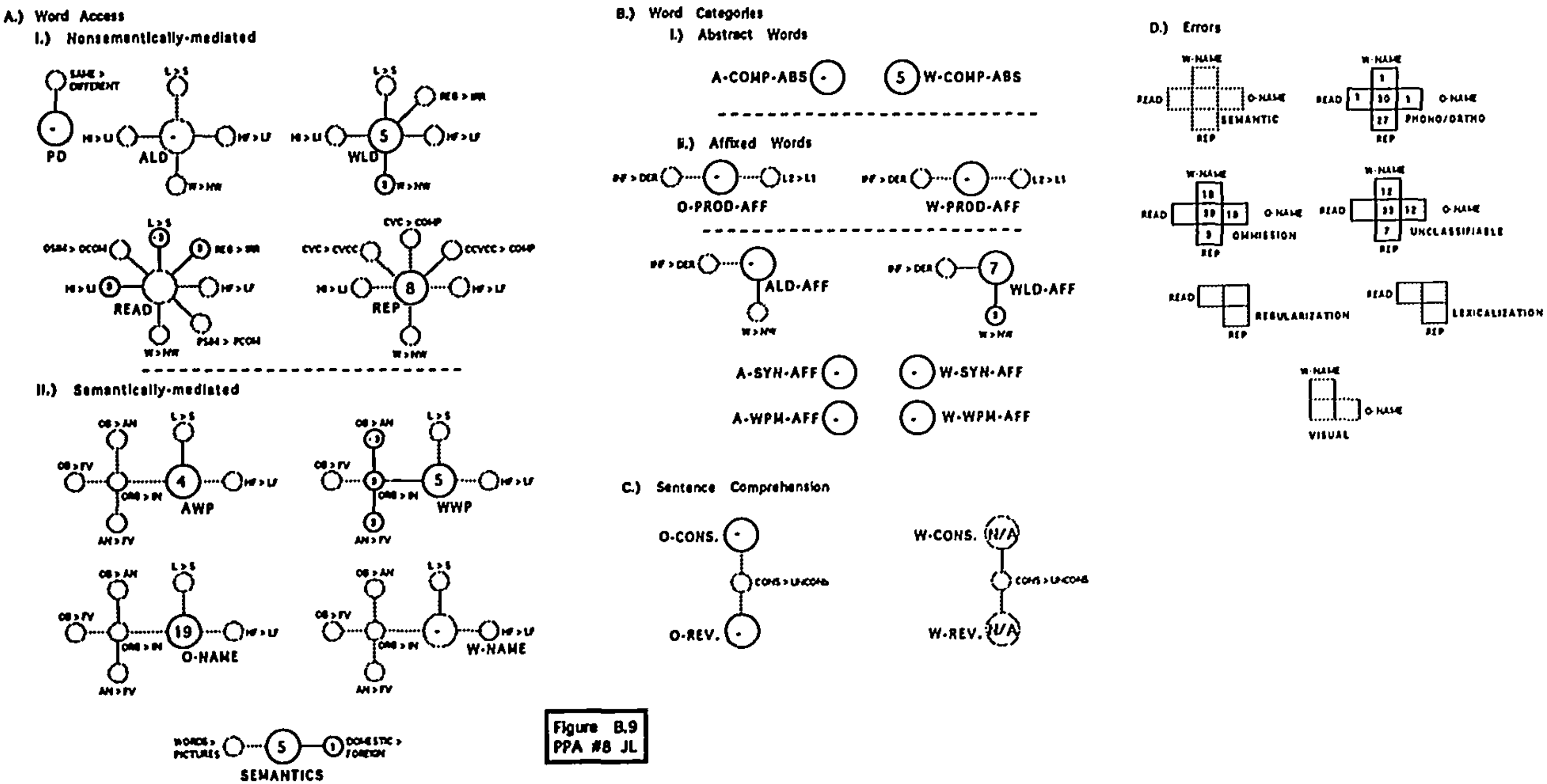


Figure B.9
PPA #8 JL

PAL Results

As illustrated in Figure B.9, JL's performance was markedly below norms on every test he completed in the battery, with the sole exception of reading aloud. He was unable to complete tests of the auditory lexical decision, auditory abstract word comprehension test, any of the affixed word tests, written naming, or the sentence comprehension tests, in each case because he either was unable to understand what was required, because he was clearly parsing the pictures incorrectly, or because he was unable to maintain the necessary concentration. Administration of the phoneme discrimination subtest was halted midway due to his apparent inability to distinguish any of the phonemes. He asked not to complete any more of the tests in the semantics battery after he had completed only two of them (28 items in all), so the results on that battery- which suggest that he has great difficulty accessing semantic information- must be considered tentative.

JL's performance is almost uninterpretable. However, it is remarkable to note that, despite his massive impairment, JL is one of only two patients in the PPA group to attain a score in the normal range (96%) on reading aloud. In light of his excellent performance at reading both words and nonwords, his very low score on repetition (37%) is equally noteworthy, indicating clearly that there is some massive interference in converting from auditory input to phonology which is not affecting the conversion of orthographic input to phonology.

Although there is not enough information to justify the inference with any certainty, it is possible that JL has the same interference with auditory comprehension as has been shown clearly in some of the other PPA patients. This would explain his apparently total inability to complete the phoneme discrimination or auditory lexical decision subtests, tests which require only minimal resources from the linguistic system since they do not tap semantic knowledge at all.

Subject 9: CM

CM is a 50 year old college-educated right-handed unilingually-English male who was first seen following a decline in cognitive abilities which had taken place over several years. A neuropsychological assessment was considered incompatible with DAT, and compatible with a dementing disorder affecting the frontal lobes and the post-central left hemisphere. The case is not a clean case of progressive aphasia, since, along with evidence of a number of linguistic deficits, there was some evidence of extra-linguistic cognitive deficits in learning and memory.

CM's presenting problems were described in some detail since he was given a neuropsychological symptom checklist, which a family member filled out. The extra-linguistic symptoms included difficulty making decisions, difficulty concentrating, memory deficits, difficulty handling multiple simultaneous tasks, difficulty with arithmetic, and sleeping more than he used to. The linguistic problems included word-finding difficulty, and difficulty in reading, writing, spelling, and articulation.

His Folstein Mini-Mental score was 16/30, indicating mild dementia. His recall of the Wechsler paragraphs and his learning curve for paired verbal associates were both profoundly impaired. Recognition memory was better, although still impaired, for verbal stimuli, but was normal for nonverbal stimuli.

His score on the Yesavage Geriatric Depression Scale indicated mild depression.

Brain Scans

A SPECT scan found global hypoperfusion, with the most pronounced attenuation in the left temporal and left frontal regions.

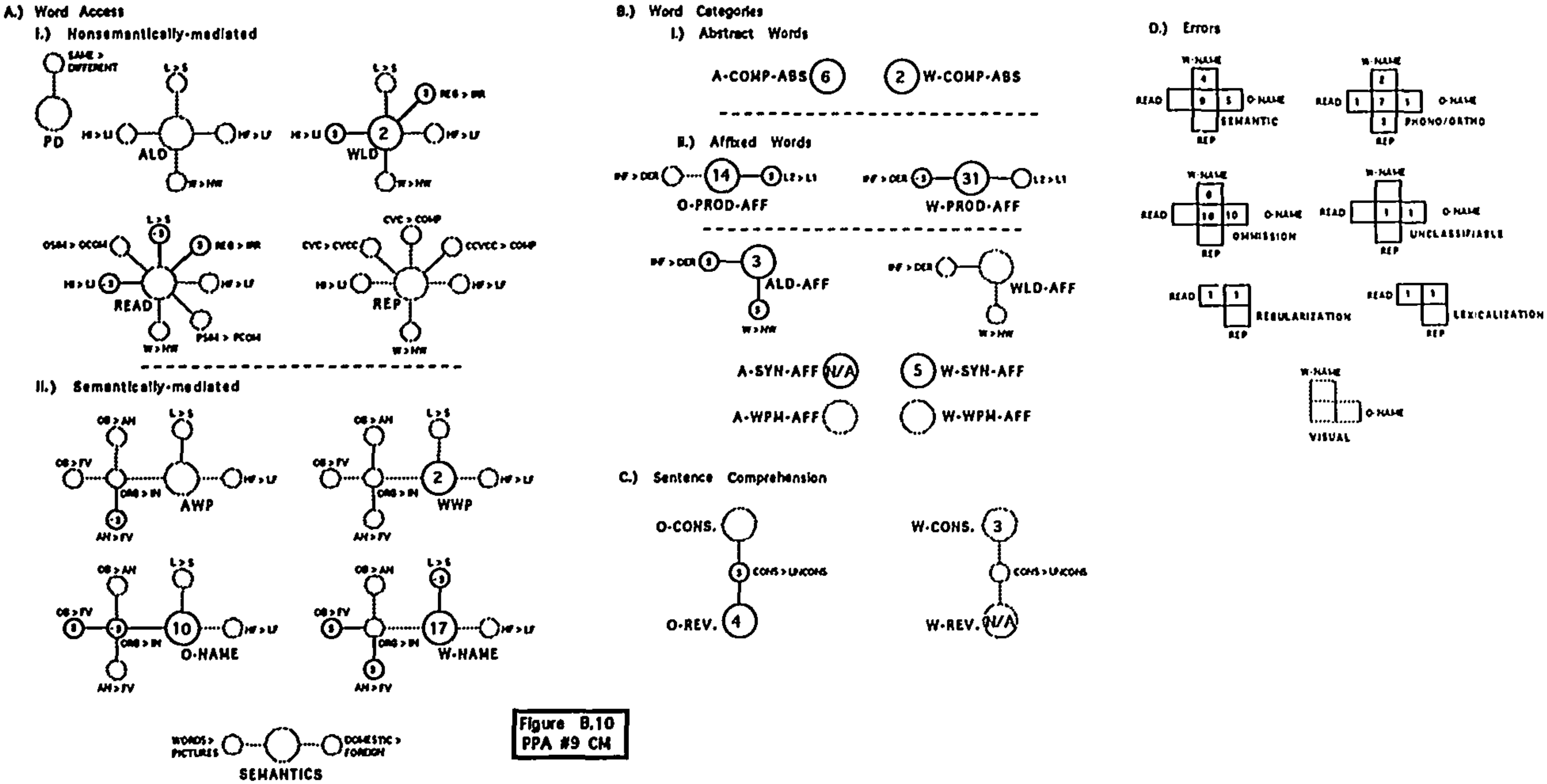


Figure B.10
PPA #9 CM

PAL Battery Results

With exception of a very mild deficit in written lexical decision, CM's performance on the five tests of nonsemantically-mediated word access was within the normal range. His deficits (as summarized in Figure B.10) include the following:

i.) Agnosia

CM has a deficit with semantically-mediated word production. Although he is able to read, repeat, and recognize words, and match words to pictures in both modalities, he has great difficulty on the subtests of oral (47%) and written (62%) naming, and of oral (17%) and written (7%) production of derived words. Although most of his errors on the naming tests were omissions, he made a relatively large proportion (oral: 5/15 written: 4/12) of semantic errors: i.e. 'antler' for 'moose', 'nipers' [nippers] for 'pliers', and 'cow' for 'ox'. DM has good access to the semantic information which is required by the tests of the semantics battery, suggesting that his naming problem may be a word-finding problem rather than a semantic access problem.

However, both the semantic nature of his errors and the modulation of his naming deficit by category casts some doubt on this interpretation, suggesting that his naming problem may reflect some difficulty accessing semantic information as well. In both modalities, CM had significantly more difficulty naming fruits and vegetables (oral: 38%; written: 38%), than he did naming objects (oral: 63%; written: 69%). In the written modality only, he was also significantly worse at naming fruits and vegetables than animals (75%). This is suggestive evidence of a category-specific anomia confined to fruits and vegetables.

ii.) Abstract Word Comprehension Deficit

CM has difficulty with comprehension of abstract words in both the oral (70%) and written (80%) modalities.

iii.) Affixed Word Production Deficit (Equivocal)

There is some uncertain evidence which suggest that CM may have specific deficit in production of affixed word forms. Although it is difficult to separate his deficit in producing affixed word forms from his more general word finding problem, especially since many of his errors on the affixed word tests were omission errors, two pieces of evidence suggest that there are two separate deficits. One is that his scores in both oral (17%) and written (7%) affixed word production (and the standardized scores for those tests) are much lower than the analogous scores in the naming tests. However, this may reflect merely the higher cognitive load required on the affixed word tests. The second piece of evidence is that CM did produce a few clear morphological errors: for example, he produced 'flied' (which he knew was wrong, but was unable to correct) for 'flew', and produced 'benefit for me' instead of 'beneficial'. Most of his written errors consisted of simply copying down the cue word, rendering their theoretical interpretation equivocal.

iv.) Sentence Comprehension Deficit

CM has a clear deficit with syntactical structure. He was at chance (60%; $\chi^2(1) = 0.1$; $P > 0.05$) at comprehension of oral semantically-reversible sentences, but significantly better (and, at 95%, nearly perfect) at comprehending the oral semantically-constrained sentences. This dissociation, the only significant dissociation between sentence types which was seen among the PPA patients, implies that the problem lies in comprehension of the syntactic structures. The deficit is clear in DM's sentence productions, which was very poor in both the oral (16%) and written (8%) modalities. Many of the errors were clearly syntactical. For example, he orally produced such sentences as 'The lady give the rattle for the baby', 'The ball to the toy to the man' and 'The parcel to the girl given to the man'. His written productions included 'Person in car on hat', 'The bag is giving the person for food' and, most interestingly of all, 'Through the ball too the father'.

Subject 10: ES

ES is a 79 year old woman of German extraction with a high school education. She has a 5 year history of progressive deterioration in her word-finding ability. On the symptom checklist, she endorsed items which include difficulty making decisions, problems with balance, increased distractability, changes in vision, increased need for sleep, decreased math abilities and driving skills, and impaired reading, writing and spelling. She shows some mild signs of dementia, as well as some behavioural signs which suggest pre-frontal involvement. Speech production is laboured and difficult. When ES does manage to speak, she speaks with a heavy German accent.

Performance on the Folstein Mini-Mental suggests moderate dementia. Her WAIS-R Verbal scale IQ was in the bottom percentile, although her Performance Scale IQ was significantly better, in the 30th percentile.

Her story recall was impaired, although she herself suggested that this deficit was secondary to her impaired speech production, pointing to her head and saying "I have the whole story in my head". There was some evidence that this was true: she showed no loss of material upon delayed recall. Her performance in learning a verbal list was similar: impaired at immediate recall, but without any loss of material upon delayed recall. She was above chance in recognition memory. Immediate and delayed recall of nonverbal material was impaired

Language

Performance on a measure of confrontation naming was 'significantly impaired'. However, she was able to match pictures to words. Her scores on tests of word fluency were significantly low. She showed no difficulty in reading or repetition.

Brain Scans

An MRI scan found moderate atrophy which was more marked frontally, as well as general ventricular enlargement.

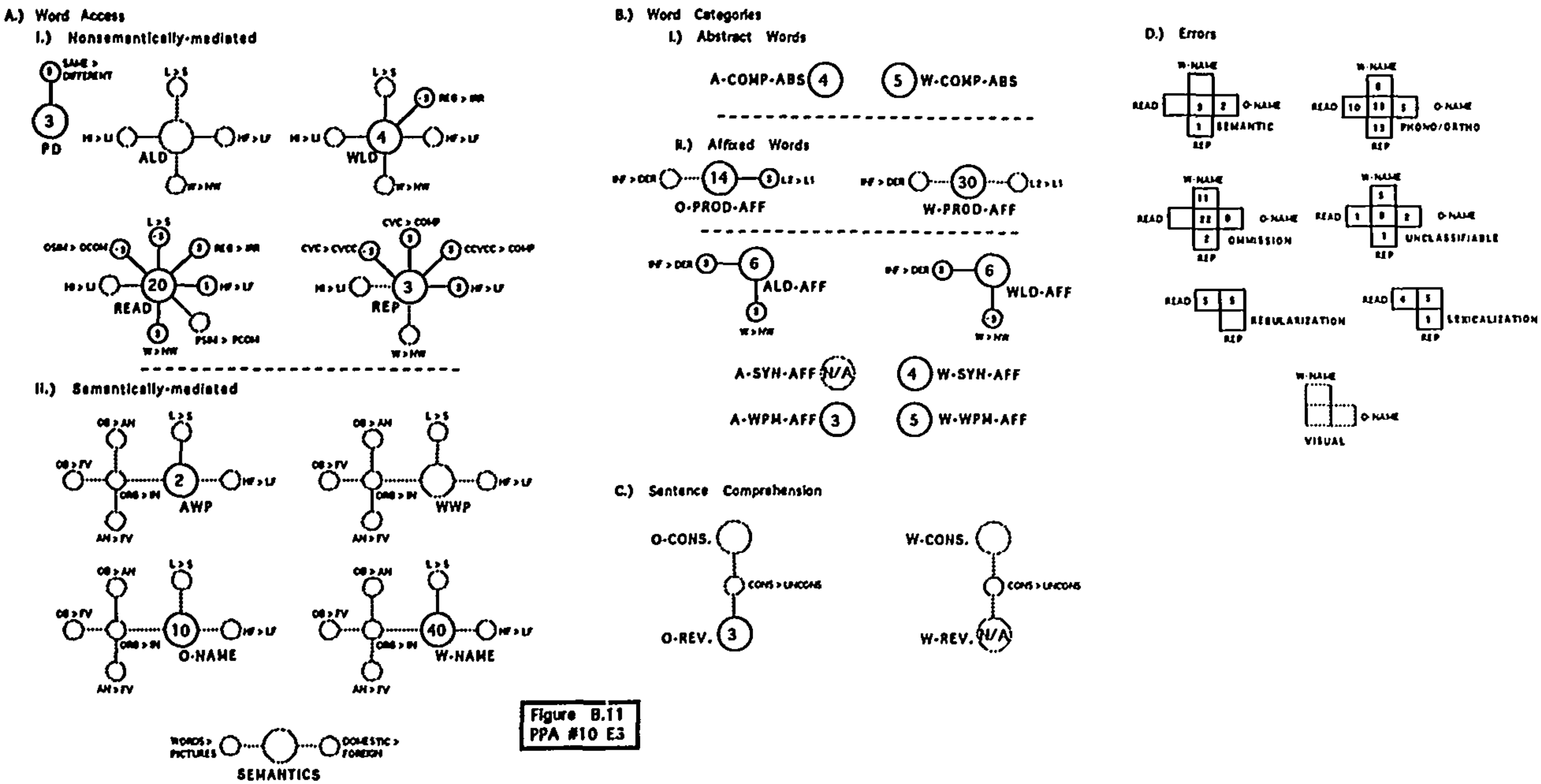


Figure B.11
PPA #10 E3

PAL Battery Results

ES's results are summarized in Figure B.11. Her major deficits are as follows:

i.) Reading Deficit

ES is severely impaired (68%) at reading. She is significantly worse at reading nonwords (60%) than words (75%), and irregular words (56%) than regular words (94%). She was also significantly worse at reading orthographically simple (short) nonwords (53%) than orthographically complex (long) nonwords (70%), and long words (81%) than short (69%) words. There was a clinically-significant frequency effect (high frequency: 88%; low frequency: 60%).

Her particular difficulty with reading both short words and short nonwords and the fact that her reading of words is modulated by frequency (which is not a characteristic which is inherent in the word itself) suggests that ES has difficulty mapping accurately into the lexicon. Further support for this suggestion is provided by her error pattern in the reading test: she made almost as many lexicalization and regularization errors (9) in reading as phonological and unclassifiable errors (11).

Her results on the repetition task (68% overall) were closely analogous, suggesting that the problem must lie in mapping into the phonological output lexicon rather than mapping into the orthographic input lexicon. On the repetition task, she is worse at repeating short (CVC) stimuli (73%) than CCVCC stimuli (93%), and also shows a significant dissociation along the frequency dimension (high frequency words: 85%; low frequency words: 50%). She does not make as many lexicalization errors in repetition of short nonwords as we might expect her to make if she had trouble mapping accurately into the output lexicon, but 50% (4 of 8) of her CVC repetition errors were analogous errors, in which she repeated short words as near-neighbour short words.

ii.) Anomia

The presumption of a deficit in mapping into the phonological output lexicon predicts that ES should show length effects on other tests which involve oral production. The only oral production test which has targets coded by frequency is the oral naming task. ES did not show a significant length effect in that test, although she showed a fairly large difference in the right direction (long: 53%; short 41%). However, the effect is muddied because ES is also anomic. She scored significantly low in both oral (47%) and written (22%) naming. She was within the normal range (89%) on a short version of the semantics battery (abbreviated to 56 items due to time constraints), and within one item of the normal range on mapping words to pictures in both the oral (88%) and written (94%) modalities. The fact that her access to semantic knowledge is intact implies that her naming deficit is not due to an agnosia.

iii.) Abstract Word Production And Comprehension Deficit

ES was significantly low on tests of both auditory (80%) and written (65%) abstract word comprehension.

iv.) Affixed Word Production And Comprehension Deficit

ES apparently has poor access to word morphology. She scored significantly below norms on every test of affixed word production and comprehension (auditory production: 17%; written production: 10%; auditory affixed word lexical decision: 73%; written lexical decision: 73%; auditory affixed word picture matching: 70%; written word picture matching: 65%; written synonym judgment: 70%).

Although the global scores on the affixed word production tests are too low to allow for an analysis of her error distribution by word type, it is noteworthy that ES was unable to correctly produce a single level 1 derived word in either modality. All of her production errors were either omission errors or true derivational errors.

In both the affixed word lexical decision tests, EM was significantly better at recognizing inflected (auditory: 88%; written: 94%) than derived (auditory: 66%; written: 63%) words.

v.) Sentence Comprehension Deficit

ES has difficulty with comprehension of auditory reversible sentences (70%). She was unable to complete either of the sentence production tests. It is difficult to draw any firm conclusions about her ability to process syntactic structure from this single low result.

Subject 11: MW

MW is a 66 year old English-speaking male with a high school education who presented three years ago with word finding difficulty and trouble understanding speech. He lived alone, making and coming to his own appointments. His speech was fluent and appropriate, and, rather remarkably, did not betray his profound deficit.

MW left the country before his neuropsychological testing could be administered.

Brain Scans

MW had both a CT and a SPECT scan. The former found evidence of left perisylvian atrophy. The latter found decreased uptake in the left fronto-parietal-temporal region and in the left thalamus, with a hole in the region of the external capsule.

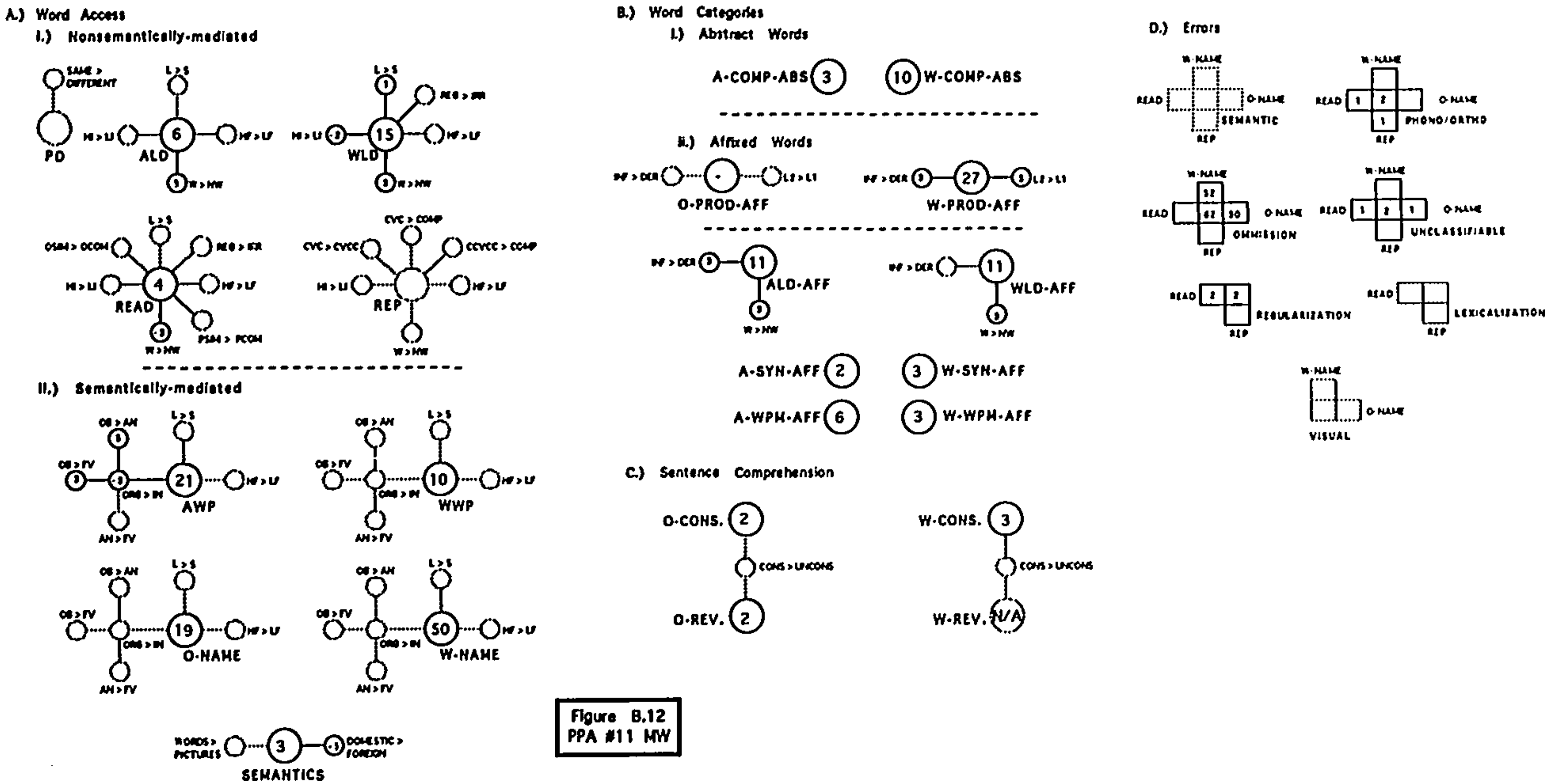


Figure B.12
PPA #11 MW

PAL Battery Results

MW scored significantly low on all but four tests in the battery: repetition, phoneme discrimination, written synonym judgments of affixed words, and at judging foreign animals only in the semantics battery (see Figure B.12). Despite this globally low performance, there were some very clear dissociations in his pattern of performance, since he is only slightly low on many tests, but almost completely incapable of answering a single question on others. The main findings are as follows:

i.) Word Recognition Deficit

MW has difficulty recognizing words in both the auditory (72%) and written (64%) modalities. He shows a consistent lexicality effect on both tests (auditory words: 95%; auditory nonwords: 33%; written words: 97%; written nonwords: 31%), which may merely indicate that he has a response bias to say yes when he is unsure. In the written modality only there was a significant effect of length (long stimuli: 78%; short stimuli: 50%), due largely to a particular difficulty with short nonwords (long words: 100%; short words: 94%; long nonwords: 56%; short nonwords: 6%).

ii.) Reading Deficit

MW has a mild reading deficit. His score on the reading subtest was high (93%), but nonetheless several standard deviations below normal due to the excellent normal performance on this test. MW made no errors all in reading nonwords, and thus scored significantly higher at reading nonwords than words.

iii.) Agnosia

MW had significant difficulty with both auditory (17%) and written (46%) word-picture matching. The written score does not differ significantly from what can be expected by chance ($\chi^2(1) = 0.6$; $P > 0.05$). The auditory score, by contrast, is remarkable in being significantly lower than what he might have been expected to attain by chance alone ($\chi^2(1) =$

28; P < 0.01). However, there is a mundane explanation for this improbable event: MW was unwilling to even hazard a guess on most of the items in the test, protesting that he absolutely no idea which choice was correct. Those items were scored wrong.

The scores on the word picture matching tests are too low to allow for meaningful analysis of the results.

MW's scores on the naming tasks are spectacularly low. He is almost completely unable to name a single common picture in either the oral (3%) or the written modality (0%).

Time did not allow for the administration of the complete semantics battery. MW's scores on a reduced set of 42 items suggest that MW has great difficulty accessing semantic information about common animals. It is difficult to interpret the apparently significant difference on his semantic battery results between his access of foreign and domestic animals without having stronger evidence that the apparent disparities reflect real differences in semantic access.

MW's globally low scores on all tests of naming and semantic access indicate that he is severely agnosic. The extent of his agnosia is clear from his performance on the oral sentence comprehension subtest. He was unable to complete this test (and was not given the written version of the subtest) because he did not know many of the words he was being asked to use in his sentences, including 'tow', 'chase' 'scratch', 'carry', 'package', 'nurse', and 'plant', among others. Clearly the difficulty with word recognition goes far beyond nouns, and might more accurately be called an 'alexia'. One of the most touching moments of the data collection period occurred when MW explained to me, with inimitable British politeness, that he was unable to produce a sentence because he did not know the word 'hug', and, what was more, he was quite sure that he had never known it! I was thus given an opportunity which may be unique in human history: to explain to a grown man 'for the first time' the concept of a hug. MW was openly delighted (as who wouldn't be?) to discover that such a thing existed, and he carefully noted the word down for future use.

iv.) Abstract Word Comprehension Deficit

MW is mildly impaired (85%) at comprehension of abstract words in the auditory modality. On the written version of the same test, he did not achieve a score which differed significantly from what could be expected by chance (30%; $\chi^2(1) = 3.2$; $P > 0.05$).

There is a relative sparing of abstract words in the auditory modality: he was 16 standard scores higher on auditory comprehension task than he was at auditory word-picture matching. The magnitude of the dissociation is misleading in this case, since MW scored lower than he needed to on word picture matching because of his refusal to guess. However, the effect is so large that it is difficult to dismiss it entirely as artefactual.

MW also showed a very small advantage (of about half a standard scores) for abstract words in the written modality. Since his scores on neither the written picture matching test nor the abstract word comprehension test were significantly better than chance, this small advantage may be confidently dismissed as artefactual.

v.) Affixed Word Production & Comprehension Deficit

MW was unable to complete the subtest of oral production of affixed words, because his did not recognize many of the words in the cue sentences. He attempted the written version, but was able to attain a global score of only 17%. He was significantly better at producing inflected words (30%) than derived words (10%), and level 2 derived words (25%) than level 1 derived words, which he was totally unable to produce. Many of his errors appear to be true morphological errors: for example, he wrote 'flyed' for 'flew', 'talket' for 'talked', and 'brighting' for 'brightness'.

Scores on all of the subtests of affixed word comprehension were significantly low (auditory affixed word lexical decision: 54%; written affixed word lexical decision: 52%; written synonym judgment of affixed words: 80%; auditory affixed word picture matching: 71%; written affixed word picture matching: 70%).

Appendix C: DAT Case Studies

In this appendix, the test results of the eleven DAT subjects involved in this study are presented. In the interests of brevity, the presentation format is simpler than the format used to present the results of the PPA subjects in Appendix B. In this appendix, there are neither justifications nor detailed scores for each deficit which is identified. Instead, a brief history and neuropsychological profile of each subject is given, followed by their A-gram, and a brief list of the main deficits identifiable from that diagram.

Subject 1: NB

NB is an 86 year old man, right handed man with a Grade 11 education.

He scored 19/30 (in the moderately impaired range) on the Folstein Mini-Mental exam, but achieved a score of 161/170 (within the normal range) on the Hierarchic Dementia Scale. His performance on the Wechsler Memory Test for immediate and delayed recall was impaired, as was his learning and recognition on the Rey Auditory-Verbal Learning Test. However, his digit span was within the normal range. On the Block Design subtest of the Wechsler Adult Intelligence Scale, he achieved an above-average score in the 63rd percentile. On the Digit Symbol test we was in the 37th percentile.

His score on the Yesavage Depression Scale was 2/30, which is not indicative of any depression.

Language

On the Boston Naming Test NB received a score of 32/60, which is in the impaired range. However, his score on the A&J Perceptual Word Picture Matching test (Chertkow, Bub, & Caplan, 1992) was perfect (30/30).

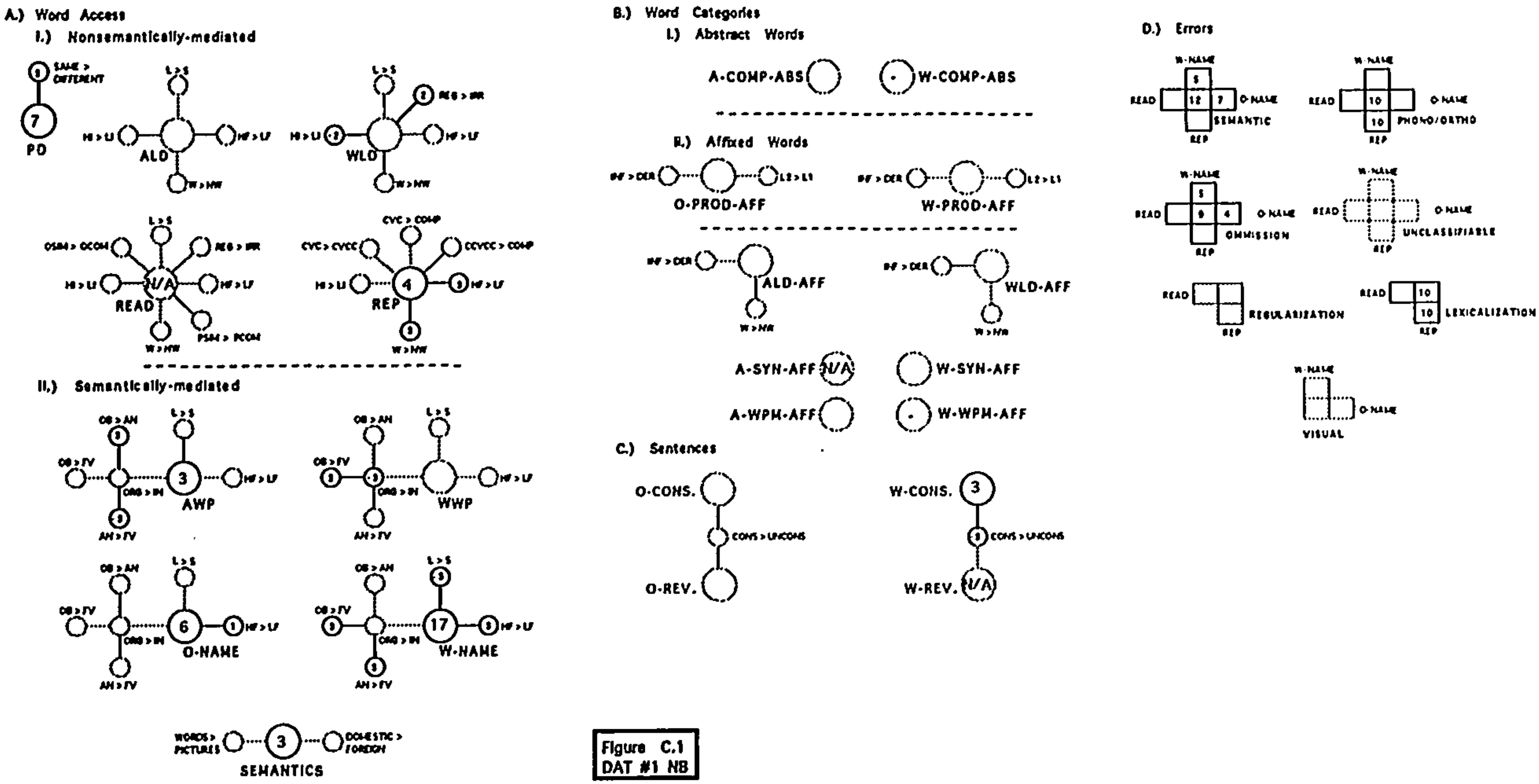


Figure C.1
DAT #1 NB

PAL Battery Results

NB's PAL results are summarized in Figure C.1.

He was unable to complete the written sentence production test. His oral sentence production was good.

Due to an administration oversight, NB's reading was not tested. However, we may infer from his scores in the normal range on both written lexical decision and written word-picture matching that his reading is intact.

NB has three main deficits, all mild:

i.) He has an auditory input deficit, tending to lexicalize nonwords.

ii.) His main problem is an agnosia. Note the preponderance of semantic errors in naming.

iii.) He has a deficit in written sentence comprehension.

Subject 2: AB

AB is an 81 year old woman, with 2 years of commercial education after high school. Although she is a native German speaker who also speaks some Czech, Ukrainian, and Hungarian, she has been living in English for many years. She has some difficulty now with self-care, relying heavily upon her husband's help in her day to day living, and presents with clear signs of dementia.

On the Folstein Mini-Mental Test, she attained a score of 19/30, in the moderately impaired range. Her score on the Hierarchic Dementia Scale was 163/180. Her memory test results show profound impairment. Her score on the Block Design subtest of the WAIS-R was average, in the 25th percentile among her age group. Her score in the Digit Symbol subtest was in the borderline range (9th percentile). Her score on the Yesavage Depression Scale (10/30) was in the average range.

Language

Rudimentary language testing revealed a profound disturbance. AB attained a score of just 18/60 on the Boston Naming Test, was impaired on the Benton Word Fluency Test, and scored significantly low (23/30) on the A&J Perceptual Word Picture Matching test.

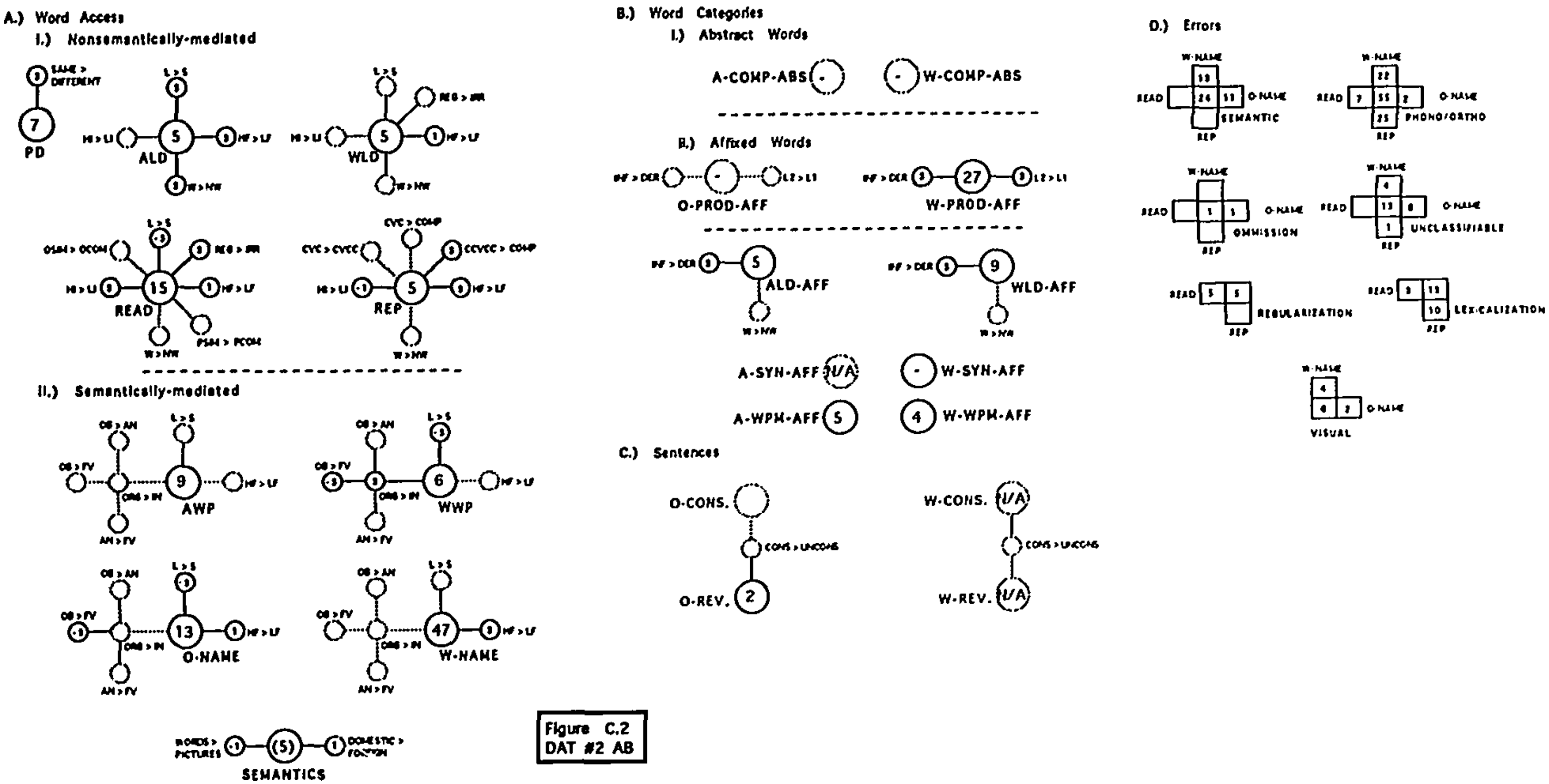


Figure C.2
DAT #2 AB

PAL Battery Results

AB's PAL results are summarized in Figure C.2.

Although she scored low on the sentence production tests because she had great difficulty following instructions, there was no evidence of a sentence level production deficit.

AB was unable to complete either of the tests requiring synonym judgments on derived words and either of the tests of abstract word comprehension, claiming the she did not recognize any of the words in these tests, or simply perseverating on her answers. As AB's husband withdrew her cooperation before testing was completed, we were not able to administer all the tests in the semantics battery. Of the two tests we did complete, AB was able to match only 1 of 14 (7%) of pictures, and 8 of 14 (57%) of words. Of the 14 Non-North American animals she saw, she was correct about only 2 (14%), while being correct on 6 of 12 (50%) North American animals.

AB was impaired on every test which she did complete, except repetition and oral comprehension of semantically constrained sentences only. Reading is very mildly impaired, mainly (but not only) for irregular words. Her written naming is particularly poor because she made both orthographic and semantic errors.

Although her performance on the oral sentence production test was poor, her low score was due to an inability to follow the instructions rather than to a syntactical deficit, as her responses were all grammatically correct.

Subject 3: DD

DD is the most highly educated of the DAT subjects, a 69-year old retired male professional with many years of university education, and a long history of intellectual interests.

His full scale in the WAIS-R was 102, with a verbal IQ of 113 and a performance IQ of 87. His immediate verbal recall was low, and delayed verbal recall very low. His nonverbal memory was slightly better but still significantly lower than the normal average.

Language

Word fluency was poor, with signs of perseveration. His score on a test of object naming was within the low normal range.

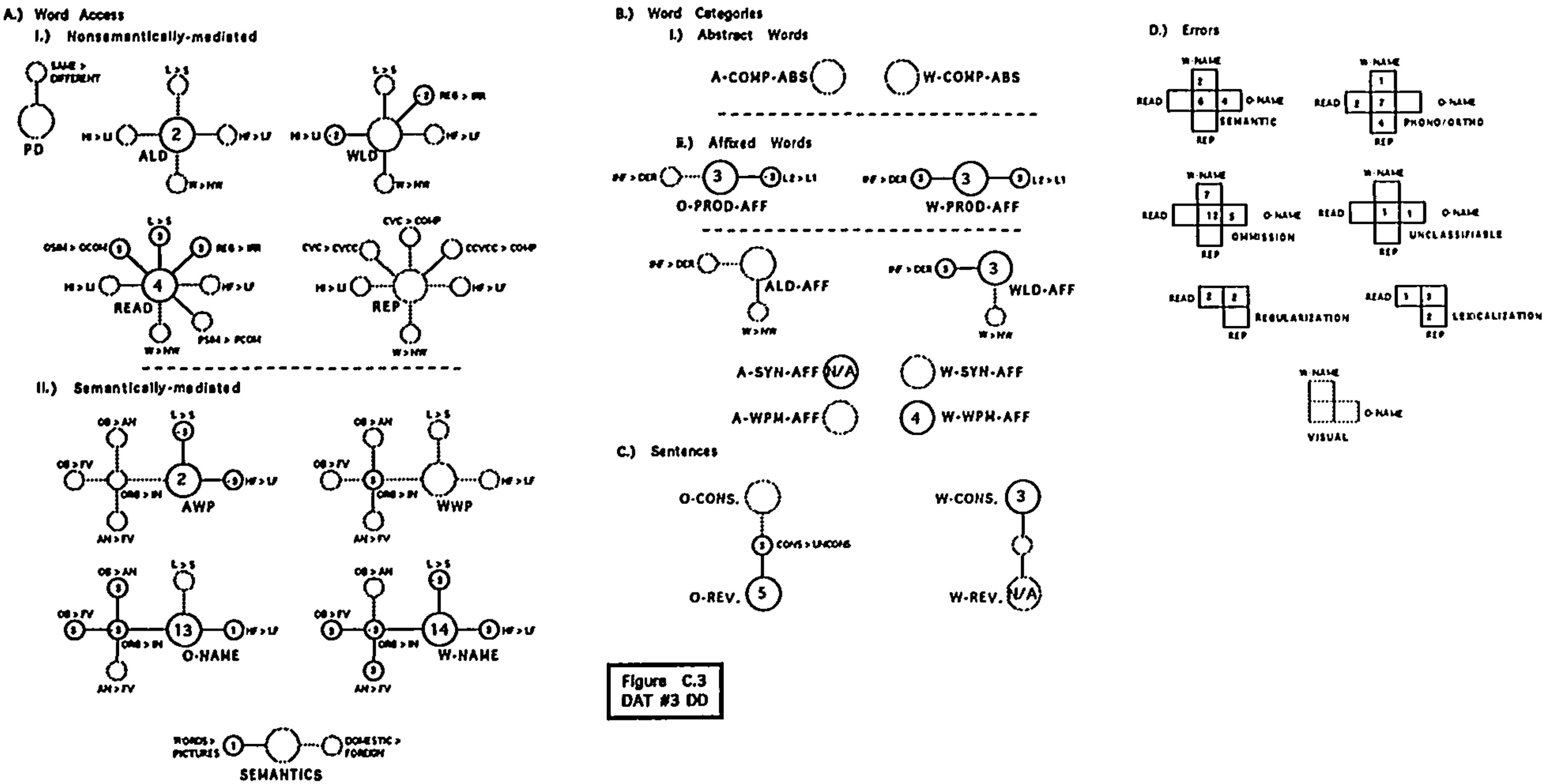


Figure C.3
DAT #3 DD

PAL Battery Results

DD's PAL results are summarized in Figure C.3.
Sentence production was unimpaired in both modalities.

DD has the following four main deficits:

i.) A mild reading deficit limited to short irregular words (surface dyslexia)

ii.) An anomia

iii.) Some sentence comprehension (but not sentence production) difficulties which are inconsistent across modalities,

iv.) Less clearly, a mild difficulty with production and written comprehension of affixed words.

Subject 4: EF

EF is a 76-year old unilingual woman (though she reported herself to be 81) with a Grade 7 education.

She was mildly impaired (20/30) on the Folstein Mini-Mental Status Test. Her verbal memory was profoundly impaired for both immediate and delayed recall. She was severely impaired (2nd percentile) on both the Digit Symbol and Block Design subtests of the WAIS-R. Her score on the Yesavage Depression Scale (14/30) suggested that she was mildly depressed.

Language

She was mildly impaired on the Boston Naming Test (52/60). Her score on both the A&J Perceptual Word Picture Matching test and the Benton Word Fluency test were in the normal range.

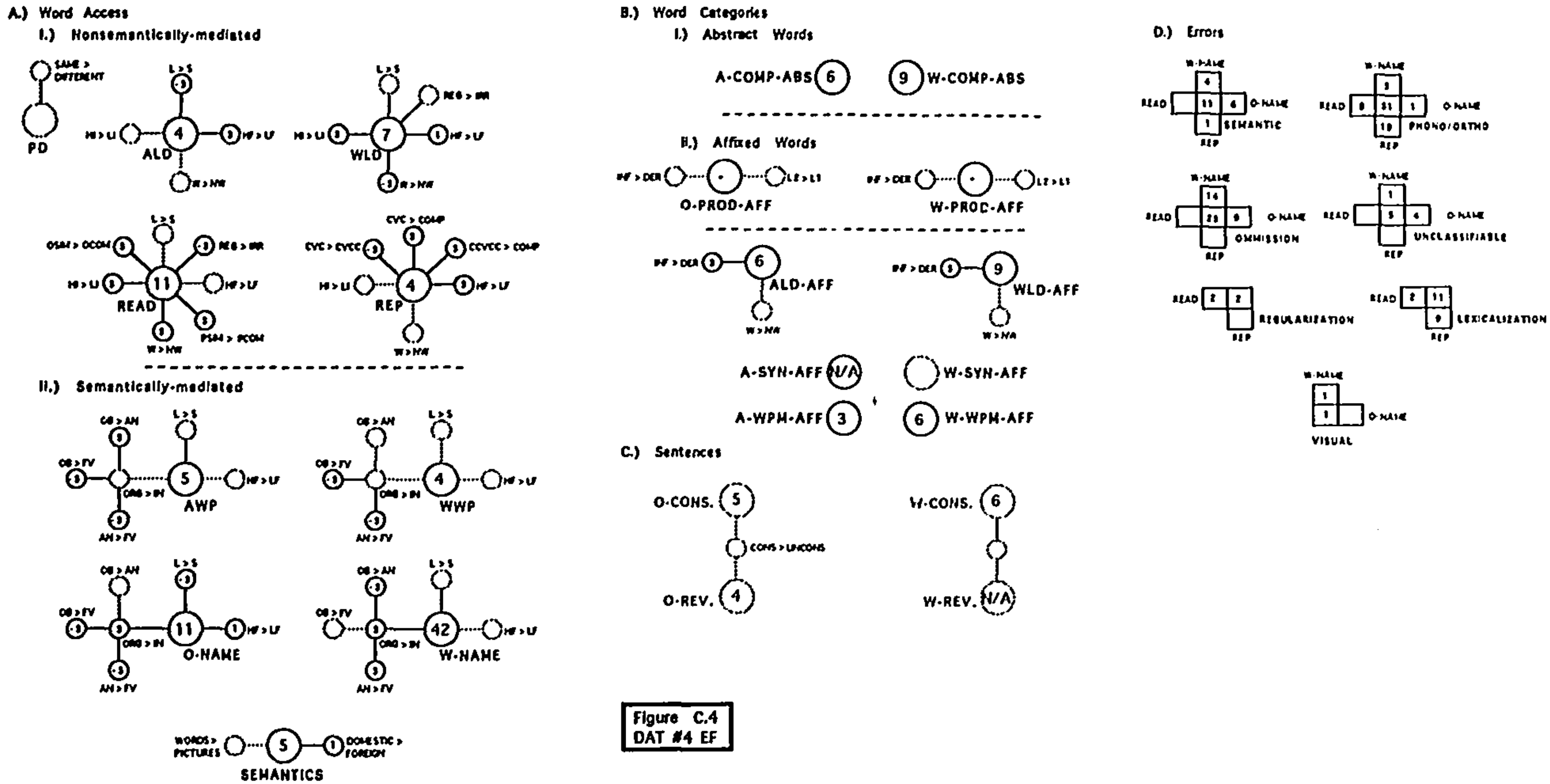


Figure C.4
DAT #4 EF

PAL Battery Results

EF's PAL results are summarized in Figure C.4.

Like AB, EF is was impaired on almost every test in the PAL battery, with the exceptions of phoneme discrimination (the lowest level test) and written synonym judgments of derived words (which has a very large normal standard deviation). She was unable to complete three tests: the two tests of derived word production, which she did not appear to understand, and the auditory synonym judgment of derived words, on which she attended only to the root. Her naming is particularly poor, especially in the written modality. There is evidence that the naming deficit is modulated by category: EF performs significantly better with stimuli from the fruits and vegetables category than with stimuli from the other two categories on six of the eight possible comparisons.

Subject 5: IK

IK is a 60 year old right handed male with 3 years of post-secondary education. He is fluent in Yiddish was well as English. He is still active in running his own business. He had presented himself three years ago with memory and word-finding problems. Testing at that time revealed impaired attention, poor insight, and impaired verbal fluency, as well as confirming his memory problems. He attained a score of 28/30 on the Folstein Mini-Mental Status test. His immediate verbal memory was in the normal range, but his delayed memory was impaired. He was in the normal range (25th percentile) on both the Block Design and Digit Symbol subtest of the WAIS-R. His score on the Beck Depression Inventory (3/63) did not suggest that he was depressed. Though eventually diagnosed as probable DAT, his symptoms are the mildest of the DAT subjects considered in this thesis.

Brain Scans

An MRI revealed enlarged sulci and dilatation of the ventricular system, which was judged to be consistent with cortical and subcortical atrophy. A PET scan showed evidence of moderate hypometabolism in the posterior parietal and occipital regions, especially on the right. The right basal ganglia were noted to be 'less prominent'.

Language

He scored perfectly on the Boston Naming Test (60/60) and on the A&J Perceptual Word Picture Matching test (30/30). His score on the Benton Word Fluency test was in the impaired range.

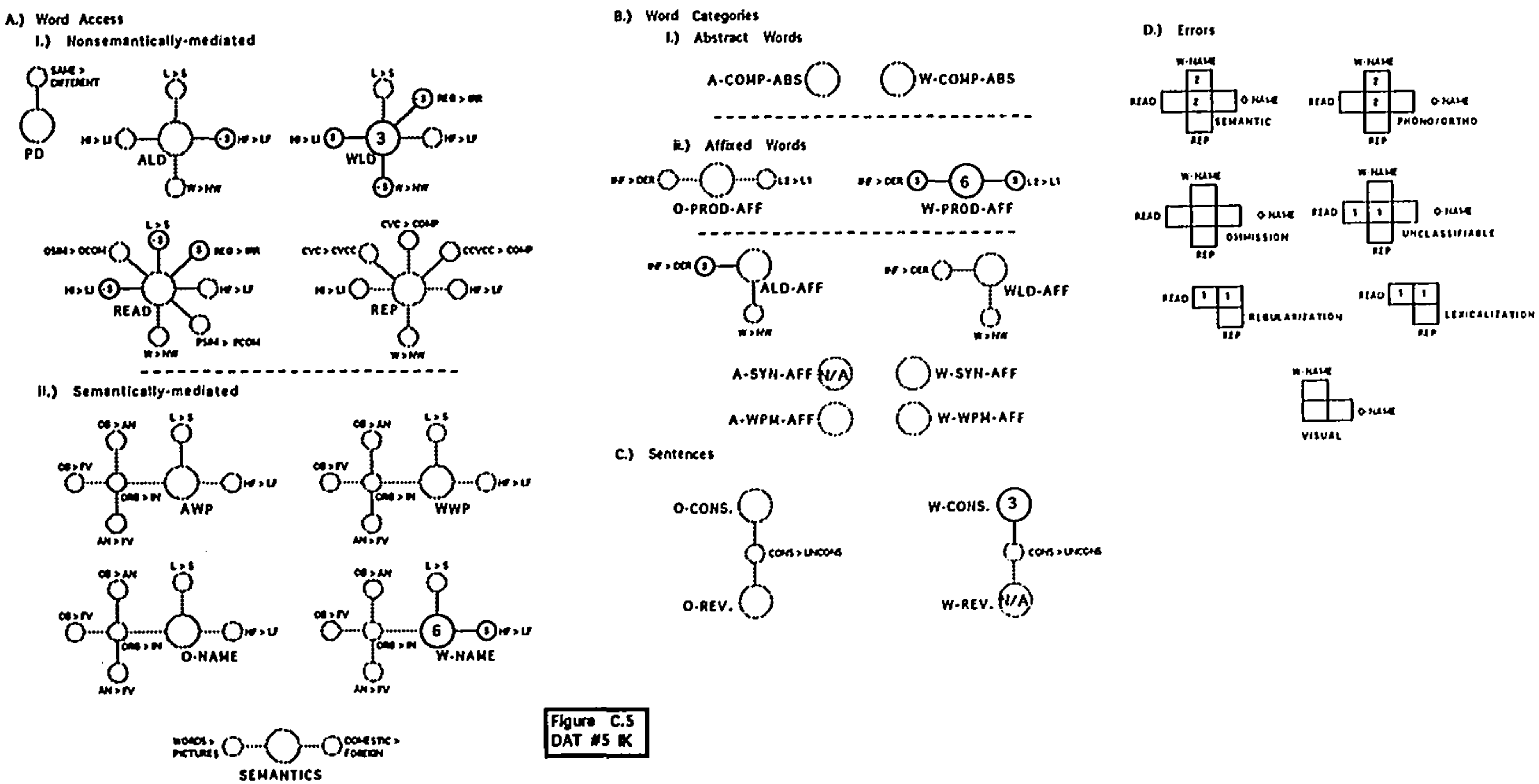


Figure C.5
DAT #5 IK

PAL Battery Results

IK's PAL results are summarized in Figure C.5.

He was unable to complete the written sentence production test, but did well on the oral production test.

IK is of particular interest since he is the highest functioning and the mildest clinically of the eleven DAT subjects in this series. His language deficits are very mild, but show some coherence inasmuch as they all relate to the written modality, with scores on the equivalent tests in the oral modality all within the normal range. He scored within the normal range on all of the PAL tests except the following four tests in the written modality: lexical decision, picture naming, production of derived words, and constrained sentence comprehension.

Both his naming deficit and his deficit in production of derived words are clearly related to a difficulty spelling, rather than to a naming or morphological processing deficit.

Subject 6: DO

DO is a 78 year old right handed man, with little formal education. Though he claimed, rather implausibly, to have left school at the age of 4, his medical chart reports that he in fact had five years of formal education. He speaks some French and Italian as well as English. He was an active boxer in his youth.

He was mildly impaired (21/30) on the Folstein Mini-Mental Status Test. He was almost totally unable to remember anything in tests of immediate and delayed verbal memory. He scored in impaired range (2nd percentile) on both the Block Design and Digit Symbol subtests of the WAIS-R. He score on the Yesavage Depression Test was low (1/30).

Language

DO was severely impaired on the Boston Naming test (24/60) but attained a perfect score (30/30) on the A&J Perceptual Word Picture Matching test. His score on the Benton Word Fluency test put him in the severely impaired range.

Brain Scans

A CAT scan showed diffuse cortical atrophy, with bi-lateral dilation of the lateral ventricles.

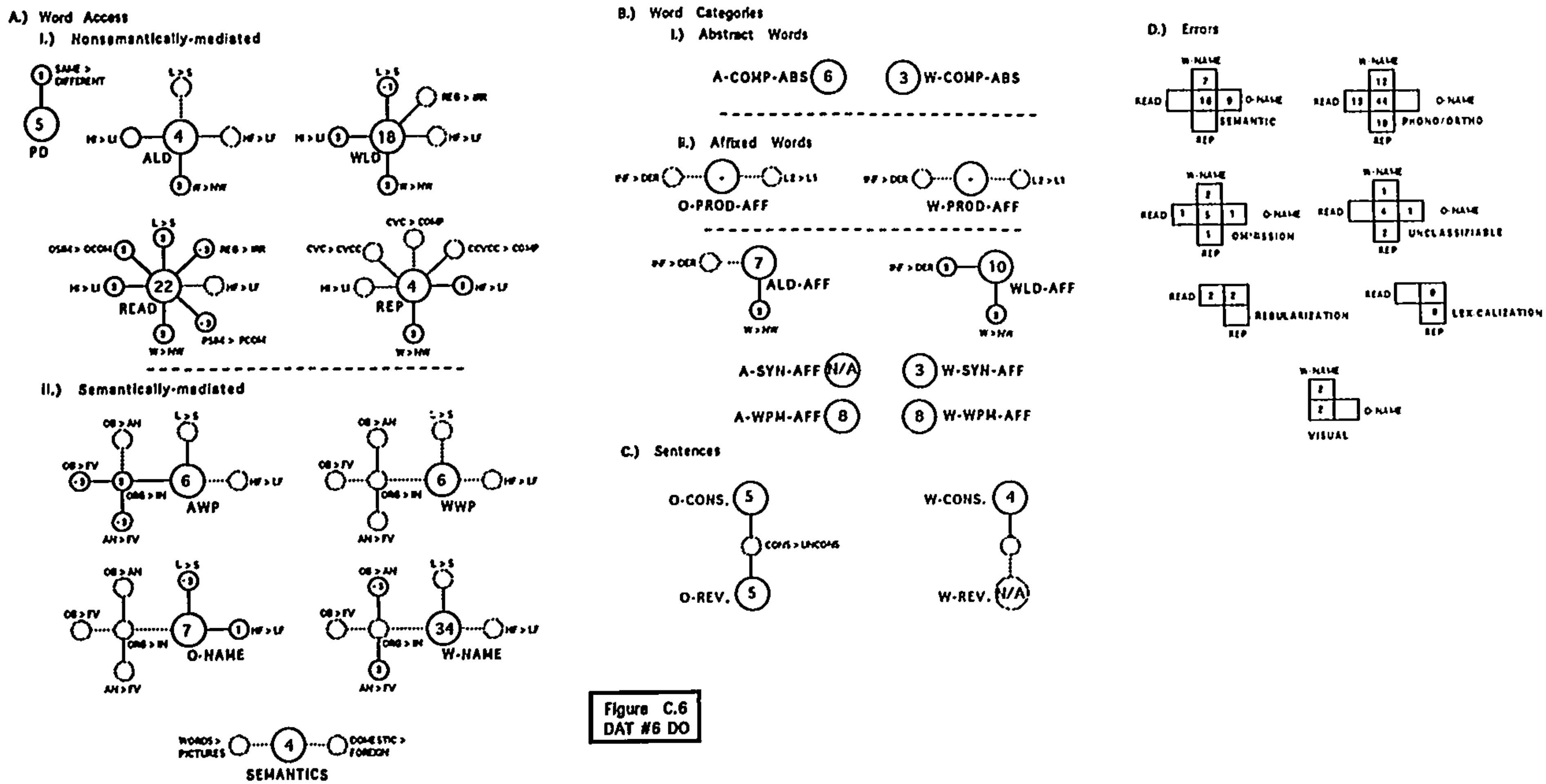


Figure C.6
DAT #6 DO

PAL Battery Results

DO's PAL results are summarized in Figure C.6.

DO scored significantly (and in most cases extremely far) below norms on every single test in the PAL battery, except for the two tests of derived word production, which he was not able to understand well enough to complete. He also scored significantly low on every global measure of the semantics battery. He was not able to complete the written sentence production test. On the oral test indicate that he was impaired.

He was remarkably consistent in being significantly better at processing words than nonwords on all six tests (lexical decision and lexical decision of derived words in both modalities, reading, and repetition) for which the distinction is relevant, a finding which is particularly notable since such perfect consistency was not seen for any other dissociation among any of the subjects in this study. As Howard and Franklin's model (Figure 2.7) makes clear, there is no theoretical reason why such a dissociation should be cross-modal. Under that model, there is no simple way to lesion the word-processing system to affect all processing of nonwords, nor does the model provide any mechanism for a 'global' modulating factor. The effect may be secondary to DO's particularly low level of education. However, the model does also offer an alternative explanation for why processing nonwords is more difficult than processing words, which in turn explains why we might expect to see a subject like DO who has particular trouble with nonwords: because there are fewer routes which can be used to process nonwords than words. Not only are more lesions required to seriously disrupt word processing than are required to disrupt nonword processing, but there is furthermore, of course, no lesion which could leave nonword processing intact while disrupting word processing. Any path which can be used to process nonwords could also be used by words. After all, words are only nonwords which mean something! DO's consistent nonword deficit does not challenge the model.

Subject 7: MR

MR is a 63 year old unilingual female.

She was mildly impaired (20/30) on the Folstein Mini-Mental Examination. Her immediate and delayed verbal memory were severely impaired. She scored in the impaired range (2nd percentile) on both the Block Design and Digit Symbol subtests of the WAIS-R. Her scores on the Yesavage Depression Scale (14/30) indicated mild depression.

Language

Her word fluency and score on the A&J Perceptual Word Picture Matching test were in the normal range. She was mildly impaired (52/60) on the Boston Naming Test.

Brain Scans

A CAT scan found mild atrophy consistent with age.

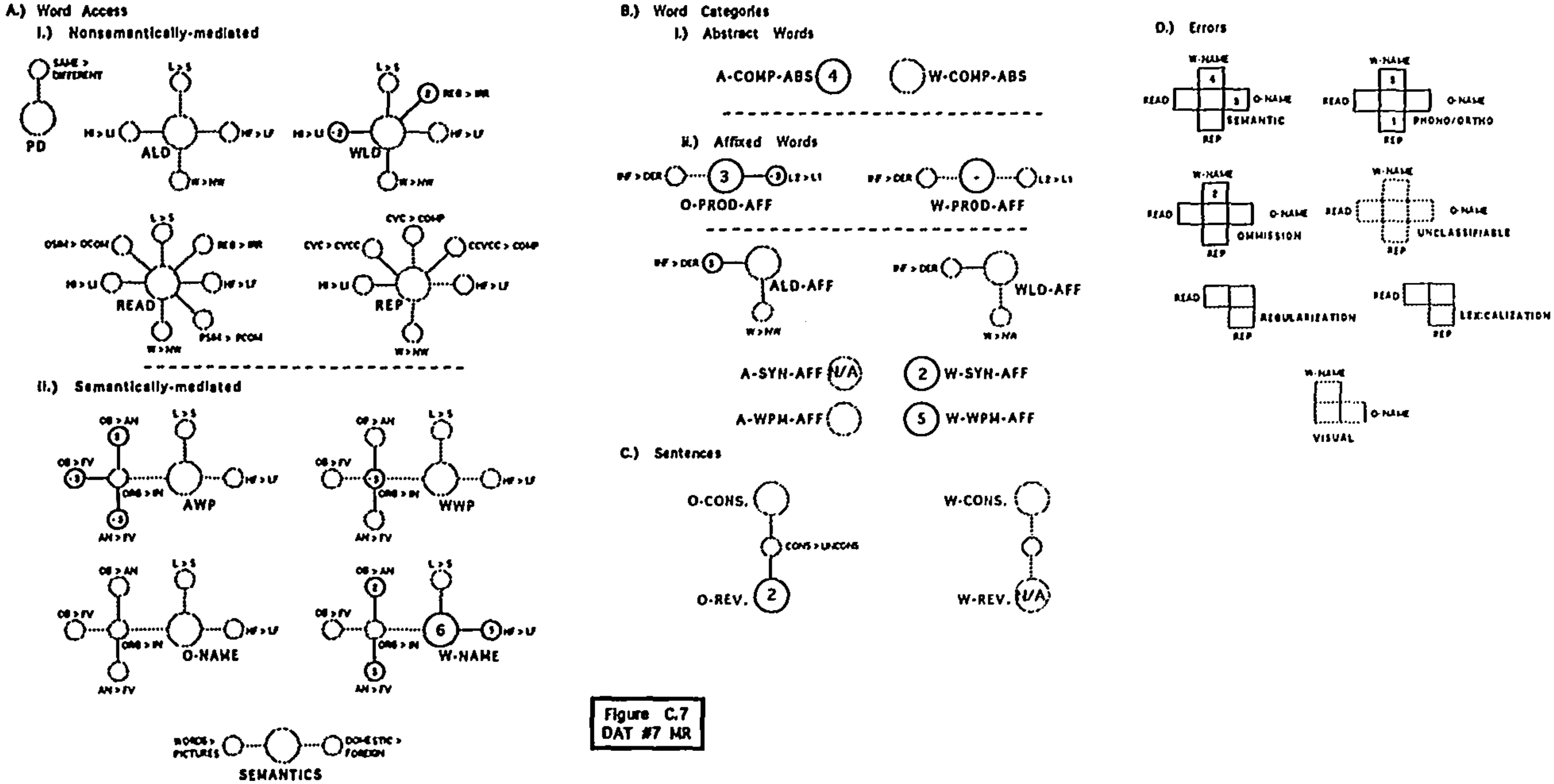


Figure C.7
DAT #7 MR

PAL Battery Results

MR's PAL results are summarized in Figure C.7.

Sentence production was unimpaired.

MR's scores from the written sentence comprehension test were lost due to a technical failure. She was unable to complete the written test of affixed word production.

The main deficits are:

i.) A mild naming deficit in the written modality, attributable to a combination of orthographic and semantic errors.

ii.) A deficit in auditory comprehension of abstract words

iii.) A difficulty with comprehension and production of derived words in the written modality only.

iv.) A difficulty with reversible, but not constrained, sentence comprehension (at least in the auditory modality) suggesting that she may be relying on pragmatic considerations rather than syntactic structure to interpret sentences.

Subject 8: DS

DS is a 79 year old unilingually-English woman who left school after Grade 11.

She scored within the normal range (28/30) on the Folstein Mini-Mental status exam and the Hierarchic Dementia Scale (185/200). Her performance on the Wechsler memory scale was impaired for delayed recall only. She was impaired on both the Digit Span and Block Design subtests of the WAIS-R. Her performance on the Benton word fluency task was within normal limits.

The score on the Yesavage Geriatric Depression Scale (6/30) did not indicate depression.

Language

DS was severely impaired on the Boston Naming Test (27/60), but scored perfectly (30/30) on the A&J Word-Picture Matching Test. She was noted to make semantic paraphasias in the course of testing.

Brains Scans

A CT scan showed moderate cerebellar and generalized cerebral atrophy. An anomaly (possibly a subarachnoid cyst, but definitely not an infarct) was noted in the left temporal lobe. It was judged by the neurologist who read the scan not to affect her dementia. A SPECT scan found decreased blod flow in the left temporal lobe only.

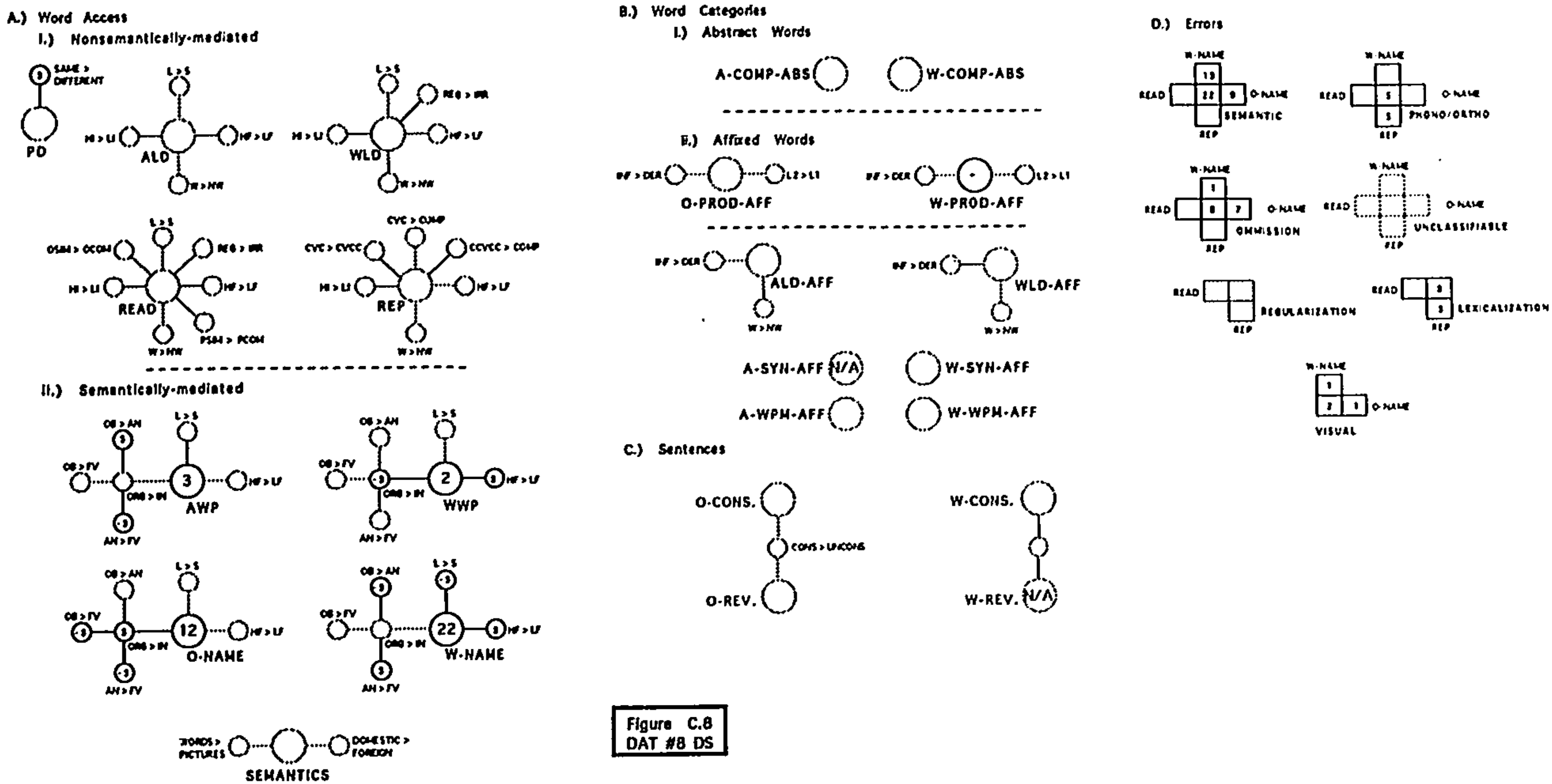


Figure C.8
DAT #8 DS

PAL Battery Results

DS's PAL results are summarized in Figure C.8.

Sentence production was unimpaired.

DS's results suggest that she has a single deficit: an anomia. The problem mapping between words and pictures is modulated by name frequency in both modalities.

Subject 9: JS

JS is an 80 year old unilingually English male, with 13 years of education. On the Folstein Mini-Mental Status Exam, he scored in the normal range (25/30). His immediate (3/25) and delayed (0/25) memory scores on the Wechsler Memory Scale and his score on the Rey Auditory Verbal Learning Test were both suggestive of severe memory impairment.

There was no indication of depression according to the Yesavage Geriatric Depression Scale (0/30).

Language

JS attained a perfect score on the Boston Naming Test (30/30) and the A&J Word Picture Matching Test (30/30).

Brain Scans

A CAT scan found evidence of mild cerebral atrophy, which was most pronounced in the left frontal lobe and left temporal horn. An MRI found moderate diffuse cerebral and subcortical atrophy.

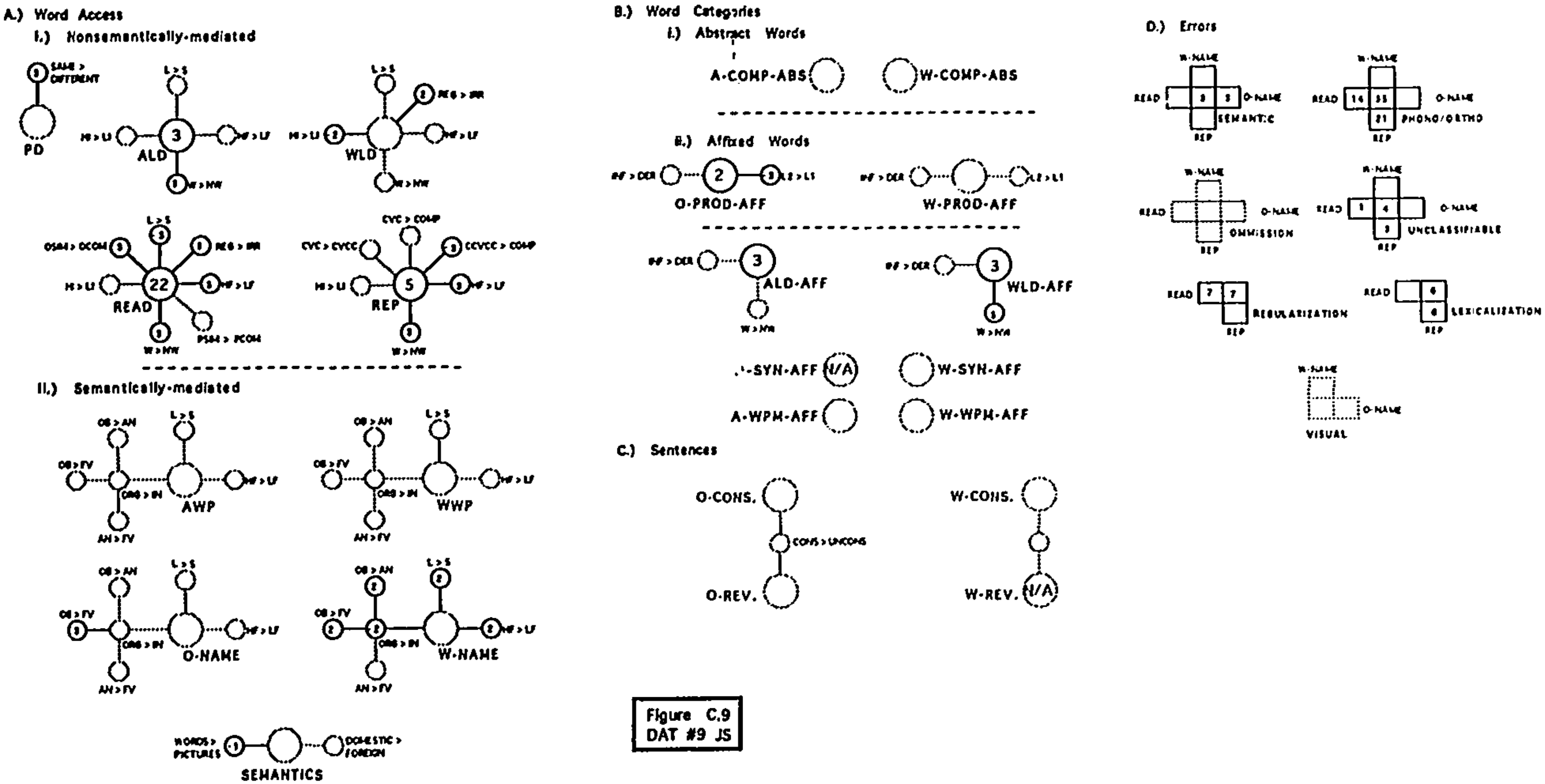


Figure C.9
DAT #9 JS

PAL Results

JS's PAL results are summarized in Figure C.9.

There was no evidence of sentence-level production errors.

JS has three deficits:

i.) A deficit in reading, especially nonwords and low frequency, irregular words (phonological dyslexia).

ii.) A deficit in repetition, especially of nonwords and low frequency words. It is interesting that this deficit mirrors his reading deficit, since most models of reading posit separate lexicons, and so do not allow for parsimonious explanations of such a consistency. JS may have a mild deficit in auditory perception, since he also attained a slightly (but not quite significantly) low score on auditory lexical decision.

iii.) An equivocal mild deficit in processing (especially comprehension) of affixed words. He scored mildly low on only two of the five derived word comprehension tests.

Subject 10: RS

RS is an 86 year old woman. She scored in the moderately impaired range (19/30) on the Folstein Mini-Mental Examination. Her verbal memory was mildly impaired. She was not depressed according to the Yesavage Depression Scale (8/30).

Language

Her performance on the Boston Naming Test (28/6) was impaired. She was not, however, significantly impaired on either the A&J Perceptual Word Picture Matching test (26/30) or the Benton Word Fluency Test.

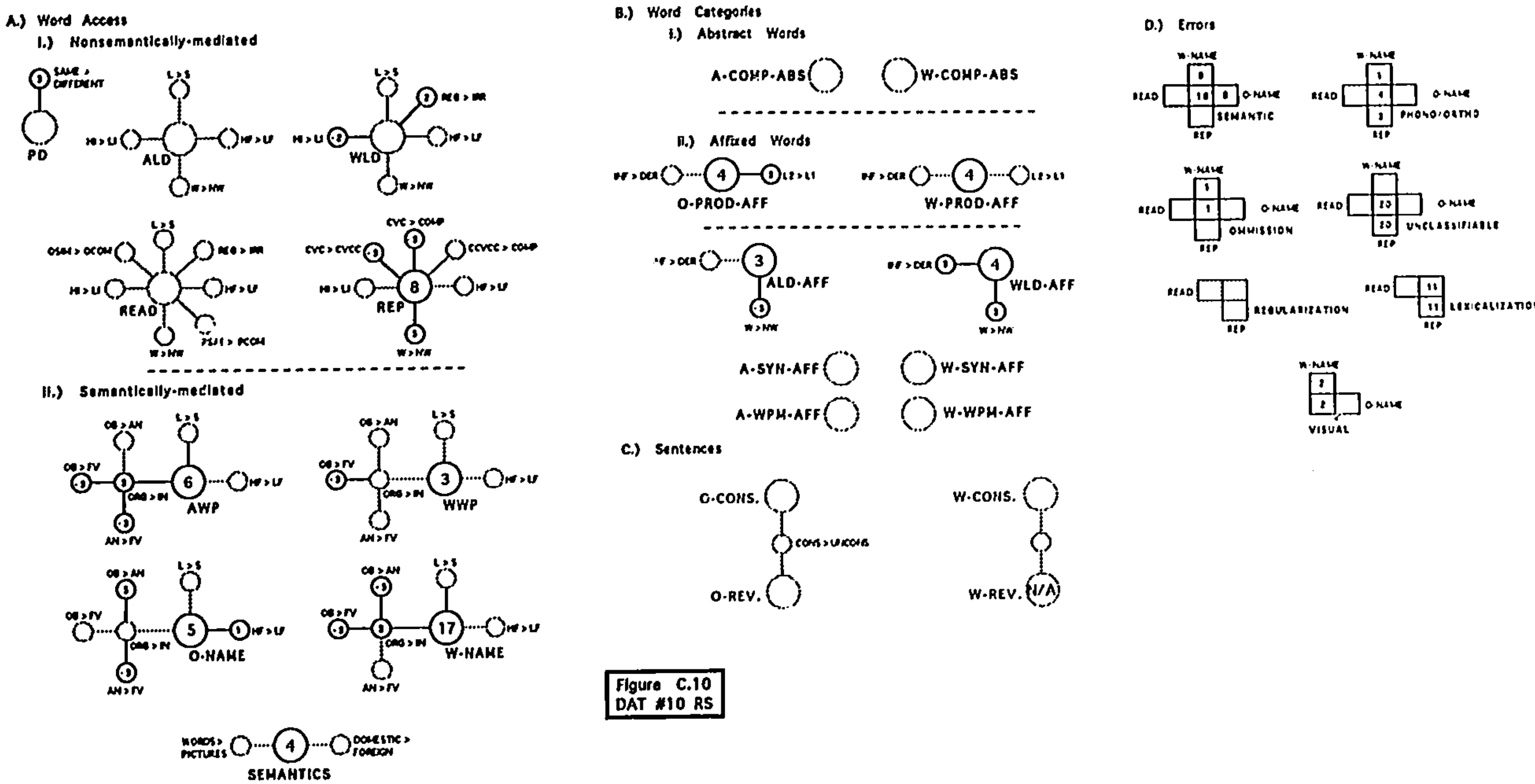


Figure C.10
DAT #10 RS

PAL Battery Results

RS's PAL results are summarized in Figure C.10.
Sentence production was unimpaired in both modalities.
RS's identified deficits are:

i.) A repetition deficit for nonwords only.

ii.) An agnosia.

iii.) An equivocal deficit in production only of affixed words. Although she scored significantly low on both tests of affixed word production, her errors do not appear to be morphological in nature.

Subject 11: YS

YS is an 82 year old right handed woman who left school after Grade 4. She speaks some Yiddish as well as English. She presented with memory problems one year before she was assessed with the PAL battery. A CAT scan at that time showed moderate cerebral atrophy, which was greater on the left than the right.

Her score on the Folstein Mini-Mental examination was in the mildly impaired range (22/30). She showed mild to moderate impairment on verbal memory testing. She scored in the borderline range (5th percentile) on the Block Design Subtest of the WAIS-R, and in the low average range of the digit symbol test. Her score on the Yesavage Depression scale (13/30) indicated that she was mildly depressed.

Language

YS was impaired on the Boston Naming Test (34/60). Her scores on the A & J Perceptual Word Picture Matching Test (30/30) and the Benton Word Fluency test were both normal.

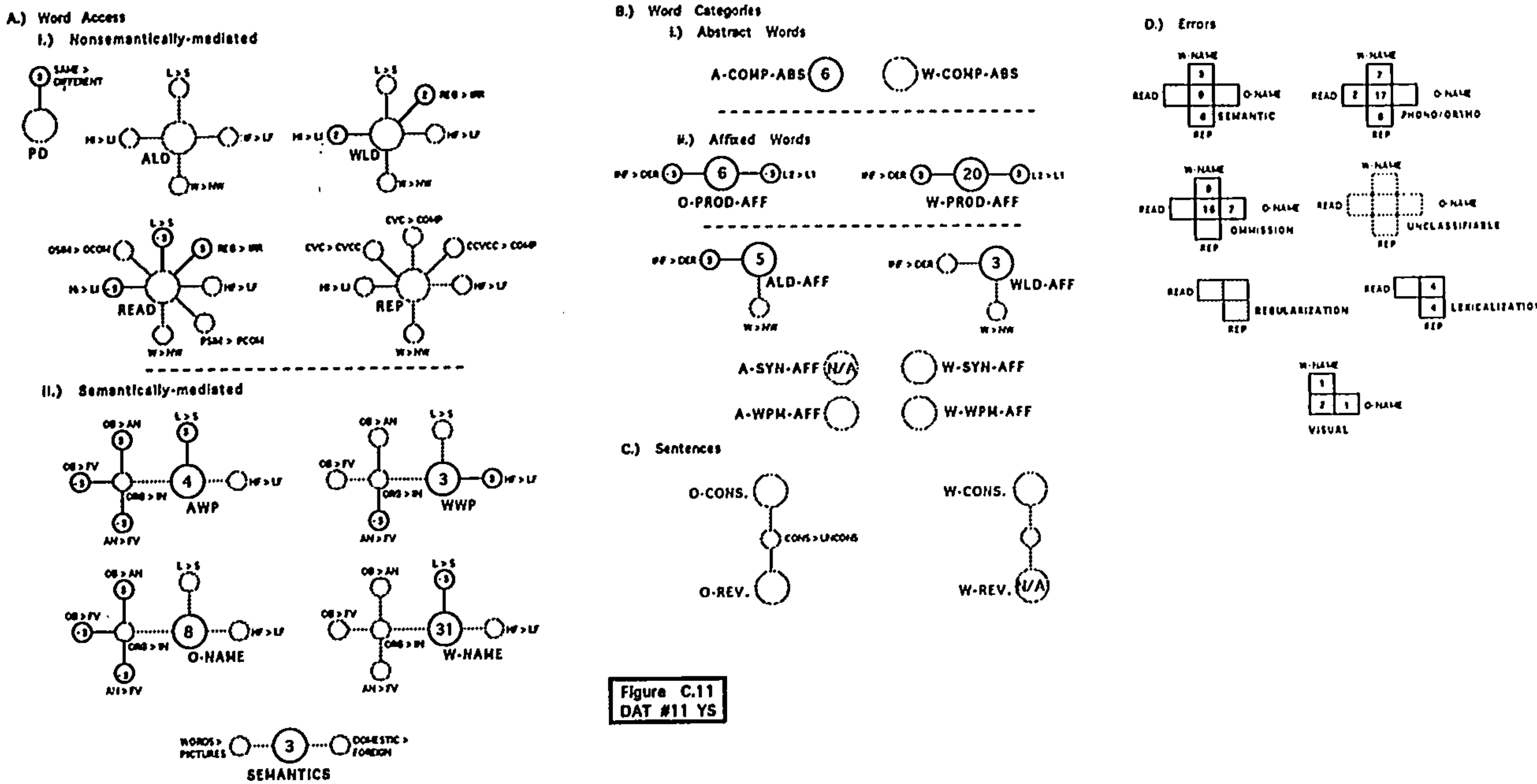


Figure C.11
DAT #11 YS

PAL Battery Results

YS's PAL results are summarized in Figure C.11.

Her sentence production tests did not indicate any difficulty with sentence production.

YS has the following language-related problems:

i.) An agnosia, which is worse for animals than for fruits and vegetables or for inorganic objects.

ii.) An equivocal deficit in both production and comprehension of affixed words. Although her scores were significantly low in both affixed word production tests, most of her errors are not unambiguously morphological in nature. Among the affixed word comprehension tests, she scored low only on the two tests of lexical decision.

BIBLIOGRAPHY

Note: Papers included in the review in Chapter 3 are marked with '•'.

Appell, J., Kertesz, A., & Fisman, M. (1982). *A Study of Language Functioning in Alzheimer Patients.* Brain and Language, 17, 73-91.

Barker, M., and Lawson, J. (1968). Nominal aphasia in dementia. *British Journal of Psychiatry, 114*, 1351 - 1356.

Barona, A., Reynolds, C. & Chastain, R. (1984). A demographically based index of pre-morbid intelligence for the WAIS-R. *Journal Of Consulting & Clinical Psychology, 52*, 885-887.

Barr, M. & Kiernan, J. (1988). *The Human Nervous System: An Anatomical Viewpoint.* (5th Ed.) Philadelphia, PA: J.B. Lipincott Company.

Basso, A., Capitani, E. & Laiacona, M. (1988) Progressive language impairment without dementia: a case with isolated category-specific semantic deficit. *Journal Of Neurology, Neurosurgery, and Psychiatry, 51*:1201- 1207

Bateson, G. (1972). *Steps To An Ecology Of Mind* New York, NY: Ballantine Books.

Bateson, G. (1979). *Mind And Nature: A Necessary Unity.* New York, NY: Bantam Books.

Bateson, G. (1991). *Sacred Unity: Further Steps To An Ecology Of Mind.* New York, NY: HarperCollins.

Bayles, K., & Tomoeda, C. (1983). Confrontation naming impairment in dementia. *Brain & Language, 19*, 98 - 114.

• Béland, R. & Ska, B. (1992). Interaction between verbal and gestural language in progressive aphasia: A longitudinal case study. *Brain And Language, 43*, 355 - 385.

• Benson, D., & Zaias, B. (1991). Progressive aphasia: A case with postmortem correlation. *Neuropsychiatry Neuropsychol Behav Neurol, 4*, 215 - 223.

Benton, A. & Hamsher, K. (1983). *Multilingual Aphasia Examination.* Iowa City: AJA Associates.

Boorstin, D. (1983). *The Discoverers: A History Of Man's Search To Know His World And Himself.* Toronto, Canada: Vintage Books.

Breedin, S.D, Saffran, E.M., & Coslett, H.B. (1994). Reversal of the concreteness effect in a patient with semantic dementia. *Cognitive Neuropsychology, 11(6)*, 617-660.

Brownston, L., Farrell, R., Kant, E. & Martin, N. (1985). *Programming Expert Systems In OPS5: An Introduction To Rule-Based Programming.* Reading, Mass.: Addison-Wesley.

Buchanan, B., & Shortliffe, E. *Rule-Based Expert Systems: The MYCIN Experiments Of The Stanford Heuristic Programming Project.* Reading, Mass.: Addison-Wesley.

Caplan, D. (1992). *Language: Structure, Processing, and Disorders.* Cambridge, MA.: MIT Press.

Caplan, D. & Bub, D. (1990). *Psycholinguistic assessment of aphasia.* Presented at American Speech and Hearing Association Conference, Seattle, Wash.

Caramazza, A. (1984). The logic of neuropsychological research and the probem of patient classification in aphasia. *Brain & Language, 21*, 9-20.

Caramazza, A. (1986). On drawing inferences about the structure of normal cognitive systems from the analysis of patterns of impaired performance: The case for single case studies. *Brain & Cognition, 5*, 41-66.

Casselli, R.; Jack, C., Petersen, R., Wahner, H., Yanagihara, T. (1992). Asymmetric cortical degenerative syndromes: Clinical and radiologic correlations. *Neurology, 42*, 1462 - 1468.

• Chawluk, J., Mesulam, M-M., Hurtig, H. et al. (1986). Slowly progressive aphasia without generalized dementia: studies with positron emission tomography. *Annals Of Neurology, 19*, 68-74.

Chertkow, H., Bub, D. & Caplan, D. (1992). Constraining Theories Of Semantic Memory Processing: Evidence From Dementia. *Cognitive Neuropsychology, 9(4)*, 327-365

• Chiacchio, L., Grossi, D., Stanzione, M. & Trojano L. (1993) Slowly progressi ve aphasia associated with surface dyslexia. *Cortex, 29(1)*, 145-52

• Cohen, L. Benoit, N., Van Eeckhout, P, Ducarne, B., & Brunet, P. (1993). Pure progressive aphemia. *Journal Of Neurology, Neurosurgery, and Psychiatry, 56*, 923 -924.

Cole, M., Dastoor, D., & Koszychi, D. (1983). The Hierarchic Dementia Scale. *Journal Of Clinical And Experimental Gerontology, 5*, 219-233.

Coltheart, M., Patterson, J. & Marshall, J. (1987). Deep Dyslexia Since 1980. In M. Colheart, K. Patterson, and J. Marshall (Ed.), *Deep dyslexia (2nd edition)* (pp. 407-451). London: Routledge and Kegan Paul.

Constantinidis, , J., Richard, J., & Ajuriaguerra, J. de (1978). *Dementias with senile plaques and neurofibrillary changes.* In A.Issacs & F. Post (Eds.), Studies in geriatric psychiatry. Toronto: Wiley.

• Craenhals, A., Raison-Van Ruymbeke, A., Rectem, D., Seron, X., & Laterre, E. (1990). Is slowly progressive aphasia actually a new clinical entity? *Aphasiology, 4:5*, 485-509.

Cummings, E.E. (1961) *Complete Poems: 1913- 1961.* Orlando, Florida: Harcourt Brace Jovanovich

Cummings, J.; Benson, F., Hill, M., & Read, S. (1985). Aphasia in dementia of the Alzheimer type. *Neurology, 35*, 394-397.

de Montaigne, M. (1946) *Essays.* (George B. Ives, Trans.). New York: The Heritage Press. (Original work published 1570)

• Delecluse, F., Andersen, A., Waldemar, G., et al. (1990). Cerebral blood flow in progressive aphasia without dementia. *Brain, 113,* 1395 - 1404.

Dick, J., Snowden, J., Northen, B., et al. (1989). Slowly progressive aphasia. *Behavioral Neurology, 2:2,* 101 - 115.

Dixon, M., Bub, D. & Arguin, M. (1995, March) *The Influence Of Semantics on Shape Identification in Category Specific Agnosia.* In the Visual Processing Symposium conducted at the meeting of the International Neuropsychological Society, Spokane, Washington.

Dostoyevsky, F. (1950). *The Brothers Karamazov.* (Constance Garnett, Trans.). New York: The Modern Library. (Original work published 1880)

Dunn, L. & Dunn, L. (1981). *Peabody Picture Vocabulary Test - Revised Manual.* Circle Pines, MN: American Guidance Service.

Emery, O. & Emery, P. (1983). Language in Senile Dementia of the Alzheimer Type. *The Psychiatric Journal of the University Of Ottawa, 8,* 4, 169 - 178.

Faber-Langendoen, K., Morris, J., Knesevich, J. et al (1988). Aphasia in Senile Dementia of the Alzheimer Type. Annals Of Neurology, 23. 4, 365- 370.

Farah, M., McMullen, P., Meyer, M. (1991). Can recognition of living things be selectively impaired? *Neuropsychologia, 29,* 2, 185-193

Feher, E., Doody, R., Whitehead, J., & Pirozzolo, F. (1990). Progressive non-fluent aphasia. [Poster abstract]. *Neurology, 40* (suppl. 1).,

Flicker, C., Ferris, S., Crook, T. & Bartus, R. (1987). Implications of Memory and Language Dysfunction in the Naming Deficit of Senile Dementia. *Brain & 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 Outpatients For The Clinician.* The Journal Of Psychiatric Research, 12, 189-198.

Fort, C. (1974). *The Complete Books Of Charles Fort.* New York, NY: Dover Publications, Inc. (Original work published 1941.)

Foster, N., Chase, T. (1983). Diffuse involvement in progressive aphasia.*Annals Of Neurology, 13,* 224-225.

Freedman, L., Selchen, D., Black, S, et al. (1991). Posterior cortical dementia with alexia: Neurobehavioural, MRI, and PET fndings. *Journal Of Neurology, Neurosurgery, & Psychiatry, 54,* 443 - 448.

Gardner, H. (1985). *The Mind's New Science.* New York, NY: Basic Books.

Gordon, B., Seines, O. (1984). Progressive aphasia 'without dementia': evidence of more widespread involvement. *Neurology, 34 (suppl. 1),* 102

• Goulding, P., Northen, B., Snowden, J., Macdermott, N. & Neary, D. (1989). Progressive aphasia with right-sided extrapyramidal signs: another manifestation of localised cerebral atrophy [letter]. *Journal of Neurology, Neurosurgery & Psychiatry, 52(1),* 128-30.

• Graff-Radford, N., Damasio, A., Hyman, B. et al. (1990). Progressive aphasia in a patient with Pick's disease: A neuropsychological, radiologic, and anatomic study. *Neurology, 40,* 620 - 626.

• Graham, K., Hodges J. & Patterson K. (1994). The relationship between comprehension and oral reading in progressive fluent aphasia. *Neuropsychologia, 32(3),* 299-316

• Green, J., Morris, J., Sandson, J., McKeel, D., Miller, J. (1990). Progressive aphasia: a precursor of global dementia? *Neurology, 40,* 423-429.

• Habib, M., Pelletier, J. & Khalil, R. (1993). Aphasie primaire (Syndrome de Mesulam). *La Presse Medicale, 22,* 757-764.

Heath, P., Kennedy, P., Kapur, N. (1983). Slowly progressive aphasia without generalized dementia. *Annals Of Neurology, 13,* 687-688.

Hier, D., Hagenlocker, K., & Shindler, A. (1985). Language Disintegration in Dementia: Effects Of Etiology and Severity. *Brain & Language, 25,* 117-135.

• Hodges, J., Patterson, P. Oxbury, S. & Funnell, E. (1992). Semantic Dementia: Progressive Fluent Aphasia With Temporal Lobe Atrophy. *Brain,* 115, 1783 - 1806.

Howard, D. & Franklin, S. (1988). *Missing The Meaning?: A Cognitive Neuropsychological Study of the Processing of Words by an Aphasic Patient.* Cambridge, Mass.: MIT Press

Irigaray, L. (1967). Approche psycholinguistique du langage des dements. *Neuropsychologia, 5,* 25 - 52.

Judson, H. (1979). *The Eighth Day Of Creation: The Makers Of The Revolution In Biology.* New York, USA: Simon & Schuster, Inc.

Kaplan, E., Goodglass, H., & Weintraub, D. (1983). *Boston Naming Test.* Philadelphia: Lea and Fibiger.

• Karbe, H., Kertesz, A., & Polk, M. (1993). Profiles of language impairment in primary progressive aphasia. *Archives Of Neurology, 32,* 193-201.

• Kempler, D., Metter, E., Riege, W., Jackson, C., Benson, D., & Hanson, W. (1990). Slowly progressive aphasia: three cases with language, memory, CT, and PET data. *Journal Of Neurology, Neurosurgery, & Psychiatry, 53,* 987-993.

Kertesz, A. & Poole, E. (1974). The aphasia quotient: The taxonomic approach to measurement of aphasic disability. *Canadian Journal Of Neurological Sciences, 1,* 7 - 16.

• Kirshner, H., Tanridag, O., Thurman, L., & Whetsell, W. (1987). Progressive aphasia without dementia: Two cases with focal spongiform degeneration. *Annals Of Neurology, 22,* 527 - 532.

Kertesz, A. (1994, June). In A. Kertesz, (Chair), *Primary Progressive Aphasia.* Symposium conducted at the meeting of the World Federation Of Neurology, Budapest, Hungary.

Konner, M. (1982). *The Tangled Wing: Biological Constraints On The Human Spirit.* New York, NY: Harper Colophon Books.

Kuhn, T. (1970). *The Structure Of Scientific Revolutions.* (2nd ed.) Chicago, Ill.: University Of Chicago Press.

Kushner, M. (1989). MRI and ^{123}I-iodoamphetamine SPECT imaging of a patient with slowly progressing aphasia. *Advances In Functional Neuroimaging, Winter,* 17-19.

• Lippa, C., Cohen, R., Smith, T., & Drachman, D. (1991). Primary progressive aphasia with focal neuronal achromasia. *Neurology, 41,* 882 - 886.

Luzzati, C. & Poeck, K. (1991). An early description of slowly progressive aphasia. *Archives Of Neurology, 48,* 228 - 229.

• Mandell, A., Alexander, M. & Carpenter, S. (1989). Creutzfeldt-Jakob disease presenting as isolated aphasia. *Neurology, 39(1),* 55-8

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

Mathews, P., Obler, L. & Albert, M. (1994). Wernicke and Alzheimer on the Language Disturbances of Dementia and Aphasia. Brain And Language, 46, 439-462.

Mattis, S. (1976). Mental Status Examination For Organic Mental Syndrome In The Elderly Patient. In L. Bellak & T.B. Karasu (Eds.) *Geriatric Psychiatry.* New York: Grune & Stratton.

McCarthy, C. & Warrington, E. (1990a). *Cognitive Neuropsychology: A Clinical Introduction.* San Diego, CA.: Academic Press, Inc.

McCarthy, C. & Warrington, E. (1990b). The Dissolution Of Semantics. *Nature, 343,* 599-599

• McDaniel, K., Wagner, M., & Greenspan, B. (1991). The role of brain single photon emission computed tomography in the diagnosis of primary progressive aphasia. *Archives Of Neurology, 48,* 1257-1260.

McWhinney, B. & Snow, C. (1992). The child language data exchange system: An update. *Journal Of Child Language. 17,* 457-472.

• Mendez, M., & Zander, B. (1991). Dementia presenting with aphasia: clinical characteristics. *Journal of Neurology, Neurosurgery & Psychiatry, 54(6),*542-5

Mehler, M., Dickson, D., Davies, P. & Hourupian, D. (1986). Primary dysphasic dementia: Clinical, pathological, and biochemical studies. [Poster abstract]. *Annals Of Neurology, 20*, 126

Merton, T. (1965). *The Way Of Chuang-Tzu* New York, NY: New Directions

• Mesulam, M-M. (1982). Slowly progressive aphasia without any generalized dementia. *Annals Of Neurology, 11*, 592-598.

Mesulam, M-M. (1987). Primary Progessive Aphasia- Differentiation from Alzheimer's Disease. *Annals Of Neurology, 22*, 533-534.

Mesulam, M-M., & Weintraub, S. (1992). Reply to diffuse involvement in progressive aphasia. *Annals Of Neurology, 13*, 225.

Monk, R. (1991). *Ludwig Wittgenstien: The Duty Of Genius*. London, England: Vintage Books.

Morton, J. (1984). Brain-based and non-brain-based models of language. In D. Caplan, A. R. Lecours, & A. Smith (Eds.), *Biological perspectives in language*. Cambridge, Mass.: MIT Press.

Murdoch, B., Chenery, H., Wilks, V. & Boyle, R. (1987). Language Disorders in Dementia of the Alzheimer Type. *Brain And Language, 31*, 122-137

Neary, D. (1990). Non Alzheimer's disease forms of cerebral atrophy. [Editorial] *Journal Of Neurology, Neurosurgery, & Psychiatry, 53*, 929 - 931.

Parkin, A. (1993). Progressive aphasia without dementia - A clinical and cognitive neuropsychological analysis. *Brain And Language, 44*, 201 - 221.

Patel, F & Satz, P. (1994). The language production system and senile dementia of Alzheim r's type: neuropathological implications. Aphasiology, 8, 1, 1- 18.

• Patterson, K., & Hodges, J. (1992). Deterioration Of Word Meaning: implications For Reading. *Neuropsychologia, 30,12*, 1025 - 1040.

Patterson, K., Graham, N. & Hodges, J. (1994). The Impact Of Semantic Memory Loss On Phonological Representations. *Journal Of Cognitive Neuroscience, 6:1*, 57 - 69.

Pinker, S. (1994) *The Language Instinct*. New York: HarperCollins Books.

Poeck, K. (1983). What Do We Mean by "Aphasic Symptoms"?: A Neurologist's View. *Brain & Language, 20*, 79-89

• Poeck, K. & Luzzati, C. (1988). Slowly progressive aphasia in three patients: The problem of accompanying neuropsychological deficit. *Brain, 111*, 151-168.

Popper, K. (1959). *The Logic Of Scientific Discovery*. New York, NY: Basic Books.

Quine, (1961). *From A Logical Point Of View*. Cambridge, Mass.: Harvard University Press

Raven, J. (1965). *Guide to using the Coloured Progressive Matrices*. London: HK Lewis.

Rey, A. (1964). *L'Examen clinique en psychologie.* Paris, France: Press Universaire De France.

Rosen, W. (1983). Neuropsychological investigation of memory, visuoconstructional, visuoperceptual, and language abilities in senile dementa of the Alzheimer type. In: R. Mayeux and W. G. Rosen (Eds.), *The Dementias*. New York: Raven Press.

• Sapin , L., Anderson, F.& Pulaski, P. (1989). Progressive aphasia without dementia: further documentation. *Annals of Neurology, 25(4)*, 411-3

• Scheltens, Ph., Hazenberg, G., Lindeboom, J., Valk, J. & Wolters, E. (1990). A case of progressive aphasia without dementia: "Temporal" Pick's disease? *Journal Of Neurology, Neurosurgery, & Psychiatry, 53,* 79 - 80.

Schwartz, M., Marin, O., & Saffran, E. (1979). Dissociations of language function in dementia: A case study. Brain And Language, 7, 277 - 306.

Schwartz, M. (1984). What the classical aphasia categories don't do for us and why. *Brain and Language, 21*, 3 - 8.

Schwartz, M. & Chawluk, J. (1990). Deterioration of language in progressive aphasia: A case study. In Schwartz, M. (ed.) *Modular Deficits in Alzheimer'Type Dementia.*. Boston, MA: MIT Press.

Shallice, T. (1988). *From neuropsychology to mental structure.* Cambridge, England: Cambrdige University Press.

Shallice, T. & McCarthy, R. (1985). Phonological Reading: from patterns of impairment to possible procedures. In *Surface Dyslexia*, K. Patterson, J.C. Marshall, and M. Coltheart (Eds.) Lawrence Erlbaum: London, England.

Sim, A. & Sussman, I. (1962). Alzheimer's Disease: Its natural history and differential diagnosis. Journal Of Nervous And Mental Disorders, 135, 489 - 499.

• Snowden, J.S., Neary, D., Mann, D.M., Goulding, P.J., Testa H.J. (1992). Progressive language disorder due to lobar atrophy. *Annals Of Neurology 31*, 174 - 183.

Steinmtez, H. & Seitz, R. (1991). Functional Anatomy Of Language Processing: Neuroimaging And The Problem Of Individual Variability. *Neuropsychologia, 29*, 12, 1149 - 1161.

• Tyrrel, P., Warrington, E., Frackowiak, R., & Rossor, M. (1990). Heterogeneity in progressive aphasia due to focal cortical atrophy: A clinical and PET study. *Brain, 113,* 1321 - 1336.

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

Warrington, E., Shallice, T. (1984). Category Specific Semantic Impairments. *Brain* , *107*, 829-854

Watzlawick, P. (ed.) (1984). *The invented reality: How do we know what we believe we know?* New York, NY: W.W. Norton.

Wechsler, D. (1981a). *Wechsler adult intelligence scale (revised).* New York: Psychological Corporation.

Wechsler, D. (1981b). *Wechsler memory scale, revised edition.* San Anonio: Psychological Corporation.

• Weintraub, S., Rubin, N., Mesulam, M-M. (1990). Primary progressive aphasia: Longitudinal course, neuropsychological profile, and language features. *Archives Of Neurology, 47,* 1329-1335.

Whitaker, H. (1976). A case of isolation of the language function. in H. Whitaker & H.A. Whitaker (Eds.), *Studies in neurolinguistics.* New York: Academic Press. Vol. 2.

Wilmer, H. (1987). *Practical Jung: Nuts And Bolts Of Jungian Psychotherapy.* Wilmette, Illinois: Chiron Publications.

• Yamamoto, H., Tanabe, H, & Kashiwagi, A. et al (1990). *Acta Neurol Scanda, 83,* 102 - 105.

Yesavage, J. & Brink, T. (1983). Development and validation of a geriatric depression scale: A preliminary report. *Journal Of Psychiatric Research, 17,* 37-49.