Neuropsychology
1992, Vol. 6, No. 1, 77–95

0894-4105/92/$3.00

An Analysis of Confrontation Naming Errors in Senile Dementia of the Alzheimer Type

EMILY LaBARGE

Department of Neurology and Neurological Surgery
Washington University

DAVID A. BALOTA
MARTHA STORANDT

Department of Psychology
Washington University

DEBORAH S. SMITH

Division of Biostatistics
Washington University

Abstract. *Confrontation naming performance, as reflected by the Boston Naming Test, was examined in 25 healthy elderly adults and 49 elderly adults with very mild (n = 25) or mild senile dementia of the Alzheimer type (SDAT). Errors were classified within 17 different categories that differentially reflected perceptual, lexical, and semantic mechanisms. The results of these analyses suggest that, early in SDAT, there is a loss of lexical information and some loss in specific semantic attributes. As the disease progresses, however, there is increasing involvement of core semantic structures. These observations were supported by analyses of linguistically related errors, four-alternative forced-choice recognition performance, and correlational analyses between performance on the Boston Naming Test and 12 standard psychometric tests.*

Keywords: Alzheimer's disease, confrontation naming, lexical processing, semantic processing.

Deficits in language function in individuals with senile dementia of the Alzheimer type (SDAT) have been well-documented (e.g., Appel, Kertesz, & Fisman, 1982; Joynt, 1984; Knesevich, Toro, Morris, & LaBarge, 1985). Such language dysfunctions include paraphasias, semantic jargon, perseveration errors, and word intrusions (e.g., Faber-Langendoen, Morris, Knesevich, LaBarge, & Berg, 1988). One of the hallmark characteristics of this language dysfunction is a difficulty in simple confrontation naming of pictures. For example, Knesevich, LaBarge, and Edwards (1986) demonstrated that the Boston Naming Test, a measure of confrontation naming, was the best predictor among a group of clinical measures of the progression of SDAT.

Address all correspondence to David A. Balota, Department of Psychology, Washington University, Box 1125, St. Louis, Missouri 63110, or via email at C39814DB@WUVMD.

Although there is general agreement that confrontation naming ability is impaired in SDAT, the source and the changes in error types as a function of the severity of the disease have not been well established (Bayles & Tomoeda, 1983). We first consider the source of the impairment and then turn to the changes in error patterns as a function of severity of the disease.

Source of the Impairment

Most current accounts of picture naming emphasize three distinct cognitive analyses that need to be completed for accurate performance: perceptual, semantic, and lexical (e.g., Snodgrass, 1984; Warren & Morton, 1982). Specifically, most theorists would agree that, when presented with a picture for naming, an individual must perceptually analyze the visual features of the picture (the perceptual component) to access the underlying conceptual representation (the semantic component). After the correct conceptual representation has been accessed, the individual needs to access the appropriate name (lexical component) for the verbal response. Hence, difficulties in picture naming performance in individuals with SDAT could be due to any or some combination of perceptual, semantic, and lexical components. Evidence in the literature points to each of these sources as the primary locus for the deficit found in confrontation naming in SDAT.

The early research in this area appeared to suggest that many of the naming errors are perceptual in nature (e.g., Lawson & Barker, 1968; Kirshner, Webb, & Kelly, 1984; Rochford, 1971). According to these investigators, demented individuals often give the name of an object that is similar in appearance to the stimulus item. More recent studies, however, have not confirmed the perceptual breakdown hypothesis (e.g., Bayles & Tomoeda, 1983; Huff, Corkin, & Growdon, 1986; Martin & Fedio, 1983; Smith, Murdoch, & Chenery, 1989; Wilson, Kaszniak, Fox, Garron, & Ratusnik, 1981). These more recent studies suggest that such perceptual errors are relatively rare in individuals with SDAT.

The semantic deficit hypothesis has received considerably more support in the literature. Here the notion is that in individuals with SDAT there is a breakdown in the semantic features or attributes that determine the core meaning of a concept. Bayles and Tomoeda (1983), Flicker, Ferris, Crook, and Bartus (1987), and Huff, Corkin, and Growdon (1986) have all suggested that the semantic component is the primary locus for the deficit in confrontation naming. In fact, Flicker et al. found that if demented individuals could not name a picture correctly, then they were also unlikely (40% correct where chance is 25% correct) to recognize the item on a four-alternative multiple-choice test. This finding suggests that the difficulty is not simply a deficit in accessing the name of the picture but that there is a breakdown in the underlying semantic information. (Of course, the fact that subjects performed significantly above chance in the multiple-choice test suggests that access to the name of the stimulus produced at least some of the performance breakdown.)

According to the lexical hypothesis, much of the deficit in confrontation naming is due to a failure to access the appropriate lexical information (i.e., there is a failure to retrieve the name of the stimulus, but there is clear access to semantics). A number of studies support the lexical deficit hypothesis. For example, low-frequency words produce more naming errors than do high-frequency words (see Barker & Lawson, 1968; Kirshner et al., 1984; Skelton-Robinson & Jones, 1984). Traditional psycholinguistic theory suggests that word frequency modulates lexical processing (e.g., Balota & Chumbley, 1984, 1985; Forster, 1978). Moreover, easing lexical access by providing the first phoneme of the target word has a substantial impact on confrontation naming errors. For example, Martin and Fedio (1983) found that performance increased by 60% when subjects were given the first phoneme of the correct answer (also see Harrold, 1988). The effectiveness of phonemic cues suggests that part of the breakdown in naming performance is indeed lexical in nature. Finally, as Nebes (1989) pointed out, the errors that individuals make in confrontation naming are often linguistically related to the target word (e.g., Bayles & Tomoeda, 1983; Smith et al., 1989). These latter results suggest that sufficient semantic information is available for the production of a related, but incorrect, response.

The present study provides further analyses of the nature of the mechanism(s) underlying confrontation naming errors in individuals with SDAT. Although the differences between perceptual and linguistic errors are rather straightforward, the differences between lexical and semantic mechanisms are much more subtle. For example, an individual may provide considerable semantic information in response to a stimulus but may be unable to retrieve the correct name. Would this pattern suggest only a lexical deficit, or is it possible that the semantic information provided is not sufficiently organized to direct the search for the correct lexical item? Hence, in the present analyses, we describe responses along differing degrees of semantic and lexical content. In addition, we provide information obtained from a four-alternative forced-choice test that provided the subject with the correct lexical item in a background of unrelated distractors. Finally, the error responses are correlated with psychometric test performance to provide converging information about the underlying mechanisms of the confrontation naming deficit.

Changes in the Characteristics of Error Patterns as a Function of Disease Progression

An important aspect of the present study is an examination of the changes in the characteristics of the error types as a function of the progression of the disease. This is an important issue because the pattern of error types may change as a function of the severity of the dementia and the neurological substrates that become involved as the disease progresses. For example, it is possible that individuals with more severe dementia may have difficulties in basic pattern-recognition processes and that only those individuals with less severe dementia produce lexical or semantic errors. Moreover, because there is some evidence that points to a lexical retrieval deficit in healthy older adults with relatively little deficit in semantic structure (e.g., Balota & Duchek, 1988, 1989; Bowles & Poon, 1985), one might expect more of a lexical breakdown early in dementia, with the semantic breakdown occurring later in the progression of the disease. In fact, differences across the studies cited earlier regarding the underlying mechanism responsible for the breakdowns in confrontation naming may be related to differences in the progression of the disease in the samples examined. There is currently considerable concern in this literature about within-group heterogeneity in the severity of the disease (e.g., Becker, Huff, Nebes, Holland, & Boller, 1988; Mayeux, Stern, & Spanton, 1985). To shed some light on this issue, we provide information about two different levels of SDAT severity, very mild dementia and mild dementia.

The Present Study

The major goal of the present study is to provide detailed information about confrontation naming performance in the Boston Naming Task as a function of the severity of the disease. There are a number of noteworthy aspects of the present enterprise. First, an attempt was made to have a clearly defined SDAT population, in contrast with some studies that have included patients of mixed etiology in their demented group (e.g., Bayles & Tomoeda, 1983). The diagnostic and exclusionary criteria developed by Berg et al. (1982; see also Berg, 1988) were used to identify subjects. Although confirmation of the diagnosis of Alzheimer's disease requires histological examination of neural tissue, it is noteworthy that 84 of the 87 individuals diagnosed clinically by this research team have indeed exhibited neurological evidence of Alzheimer's disease at autopsy.

Second, the Washington University Clinical Dementia Rating Scale (CDR; Berg, 1984, 1988; Hughes, Berg, Danziger, Cogen, Martin, 1982) was used to identify three different groups of subjects, which varied in disease severity: a control group with no dementia, a group of very mildly demented individuals, and a group of mildly demented individuals. The very mild group is of particular interest because these are individuals in whom mild cognitive impairment due to SDAT is suspected but not sufficiently severe for a diagnosis of definite dementia. However, in a recent study by Rubin, Morris, Grant, and Vendegna (1989), 11 of the 16 subjects followed for 84 months did indeed progress to a more severe form of dementia or had Alzheimer's disease diagnosed upon autopsy. Thus, the very mild stage, by and large, represents an early stage in the development of Alzheimer's disease. The importance of the inclusion of this group is that we were able to identify aspects of naming errors that are found at two relatively early yet distinct levels of the disease progression.

Third, the present study provides information about the involvement of both lexical and semantic structures. In pursuit of this later goal, data from a four-alternative multiple-choice test are provided to tap the lexical involvement, along with data from a semantic rating measure developed to identify the extent to which core semantic structures are involved in producing the errors. Finally, to identify correlated changes in cognitive performance across severity of dementia, we present data concerning the relationship between performance on the Boston Naming Test and performance on a number of standard psychometric tests.

Method

Subjects

Beginning in 1984, a sample of 74 older adults (65 to 83 years old) was selected for longitudinal study by the Memory and Aging Project at the Washington University Medical Center; 25 subjects were diagnosed as having very mild dementia, 24 subjects were diagnosed as having mild dementia, and 25 subjects served as healthy controls. The demographic data concerning age, education, and socioeconomic status (as measured by Hollingshead's [1957] Two-Factor Index of Social Position) are displayed in Table 1. Although there was a main effect of group on all of these dimensions (all $ps < .05$), these demographic variables did not substantially correlate with performance on the Boston Naming Test (rs ranged from $-.26$ to $.32$). Moreover, the differences in these demographic characteristics across groups were not systematic. That is, whereas the control group was slightly younger than the very mildly demented group, the very mildly demented group was slightly older than the mildly demented group. Moreover, there was no reliable difference between the controls and the very mildly demented individuals in either education or socioeconomic status.

The participants met the diagnostic and exclusionary criteria described by Berg (1988) and Berg et al. (1982). Those individuals with severe hypertension, stroke, psychiatric illness or other mental disorders, or who were taking medication that can produce dementia were excluded from the present sample. In addition, tests to determine reversible causes of

Table 1
Age, Years of Education, and Socioeconomic Status of Normal Control, Very Mildly Demented, and Mildly Demented Subjects

Variable	Control (*n* = 25)	Very mildly demented (*n* = 25)	Mildly demented (*n* = 24)
Age (years)			
M	71.24	75.35	73.19
SD	4.83	5.41	5.04
Range	65–83	65–83	66–83
Education (years)			
M	14.04	13.44	11.54
SD	2.50	4.19	2.90
Range	8–19	6–21	8–18
Socioeconomic status			
M	2.96	3.16	3.83
SD	0.78	0.98	0.86
Range	1–4	1–5	2–5

dementia were conducted for all demented individuals.

The diagnosis and staging of SDAT was based on a 90-min interview (with both the subject and a collateral source) and a neurologic examination, which included a check of near vision. Subjects were rated on the CDR, by experienced clinicians, as healthy (CDR = 0), very mildly demented (CDR = 0.5), or mildly demented (CDR = 1.0).[1]

The CDR is based on clinicians' evaluations of cognitive abilities in the areas of memory, orientation, judgment and problem solving, community activity, home and hobby activity, and personal care. Interviews and examinations were conducted by one of eight physicians (four neurologists and four psychiatrists). The interviews were videotaped for independent rating by one of the other seven physicians. Both reliability of the CDR and validity of the diagnosis (based on autopsy confirmation) by this research team have been excellent (Burke, Miller, Ruben, et al., 1988; Morris, McKeel, Fulling, Torack, & Berg, 1988; Morris, McKeel, Price, et al., 1988).

Procedure

Boston Naming Test. This 60-item test (Goodglass & Kaplan, 1983) was administered to all individuals as part of a larger battery of psychometric tests completed when individuals entered the longitudinal study (Botwinick, Storandt, & Berg, 1986; Storandt & Hill, 1989). The Boston Naming Test is a standard psychometric test of confrontation naming in which simple line-drawn pictures are presented to subjects for a naming response. The typical score on this measure is the number of items correctly identified, with higher scores representing better performance.

[1]We use the term *very mild* here instead of the term *questionable* because of the recent study by Rubin et al. (1989), cited earlier, indicating that the individuals at the CDR .5 level, by and large, are in an early stage of Alzheimer's disease.

All subjects viewed sequentially all 60 items of the Boston Naming Task. A 20-s answer interval was allowed for each item, and all responses (multiple for some individuals) were recorded, including phonological errors and gestural responses. If the subject did not correctly name the target picture on a given trial, then a card with a reproduction of the same picture and with four words printed underneath the drawing was presented. One of the words was the correct name of the picture, and the remaining three words were foil items that were matched to the correct word in in terms of visual printed frequency (Carroll, Davies, & Richman 1971; Thorndike & Lorge, 1944). Letters were printed in black on a white background and were 11 cm in height.

Psychometric battery. In addition to the Boston Naming Test, a psychometric test battery was given to all subjects to assess basic cognitive functioning. This battery included tests of memory, psychomotor performance, language skills, and visuospatial performance, along with four subtests from the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955). The Wechsler Memory Scale (WMS; Associate Learning subtest; Wechsler & Stone, 1973), Benton Visual Retention Test (picture memory; Benton, 1963), WMS Logical Memory subtest (surface-level story memory), and WMS Forward and Backward Digit Span tests all tap basic memory performance. The adult intelligence measures included the WAIS Information, Comprehension, Block Design, and Digit Symbol subtests. Visual perceptual motor performance was tapped by the copying version of the Benton Copy test (copying a geometric figure) and the Trail Making A (connecting dots in a numerically configured pattern; Armitage, 1946). The WMS Mental Control subtest tapped speeded performance in producing well-rehearsed sequences, such as the alphabet or counting by threes to a given number. Finally, the Word Fluency test (Thurstone & Thurstone, 1949) best reflects basic lexical retrieval processes; in this task, subjects generate as many words beginning with a given letter (*P* or *S*) as they can, for a duration of 60 s for each letter.[2]

Results

Boston Naming Test Performance

Errors on the Boston Naming Test were coded by two independent raters according to the taxonomy presented in Table 2. Similar error classification systems in confrontation naming performance have been reported recently by Albert, Heller, and Milberg (1988) for healthy older individuals and by Smith et al. (1989) for individuals with SDAT.

The 17 types of errors were divided into three general categories: no content, linguistically related, and perceptually related.[3] No-content errors are those in which no interpretation of the drawing is evident (e.g., the subject responds, "I know that, but I can't think of

[2] The sample of the present study was part of a larger sample for which Boston Naming Test total scores were reported (Storandt & Hill, 1989). Types of errors, however, were not recorded for the remainder of the larger sample; therefore they could not be included in the analyses reported here.

[3] Actually, there was one final type of error observed that is not included in these tables. These were gestural errors that were not clearly related to the target item. For example, one subject made a gesture of brushing teeth when presented with a picture of a tennis racket. These errors were very infrequent and occurred only in the mildly demented individuals (8.3%). Because it is unclear which category these errors should fall under, we decided to report this error type separately. However, inclusion of this error type in any of the reported analyses does not substantially change the observed pattern of data.

Table 2
Scoring Guidelines for Types of Errors on the Boston Naming Test

Type of error	Examples
	No content
Empty phrase	I don't know Can't think of it
No interpretation possible	No response or jargon
	Linguistically related
Phonological	
Phonologically related	Pelican = pentagon Unicorn = hornicorn Rhinoceros = nostros Sphinx = phoenix
Semantic	
Same category	Latch = hasp
Super- or subordinate	Camel = animal Asparagus = vegetable
Function	Funnel = used for pouring Compass = makes circles
Attribute	Beaver = eats wood or builds dams Volcano = fire
Context	Stethoscope = doctors use it Sphinx = found in Egypt
Description	Noose = a rope with a slip knot
Acoustic	
Meaningful sound	Whistle = make a whistling sound or blow noiselessly Volcano = make a whooshing sound
Pantomime or gesture	
Gesture	Comb = gesture to head like combing Accordion = swing arms and hands like playing

(table continues)

it" or "Can't say it"). Linguistically related errors are those that link the drawing to the response in terms of a linguistic dimension (e.g., a subject might say the word *cat* in response to the stimulus *lion*). Perceptual errors are those that identify incorrectly either part or all of the stimulus drawing (e.g., a subject might call the *harmonica* an *apartment building* when, in fact, the picture of the harmonica can be viewed as an apartment building, given a particular orientation). With respect to the distinction between perceptual, lexical, and semantic mechanisms underlying the breakdown in confrontation naming, the linguistic errors should reflect a loss in lexical (and potentially semantic) information because these errors are often lexically related to the correct target response. No-content errors should reflect a loss in semantic information because there is no evidence that there is sufficient semantic information available to produce even a related response. Of course, the perceptual mechanism should be

Table 2 (continued)

Type of error	Examples
	Perceptually related (visually)
Whole	Whistle = trailer hitch or pacemaker
	Knocker = chandelier
	Igloo = turtle or spider's web
Part Dart = feather	
	Rhinoceros = two big horns
Perspective	Harmonica = windows or apartment building or file drawers
Function	Dart = nurse to give shot
	Broom = wash my clothes
Attribute	Beaver = fella who goes underground
	Mask = a bad picture
Context	Wreath = see 'em at wedding
	Whistle = hanging on a tree limb

primarily reflected by the perceptual error types.[4]

To check the interrater reliability for the scoring of the three categories of errors, we calculated kappa coefficients for each Boston Naming Test item (no errors were made on Item 2, *tree*). If subjects made multiple errors on an item, only the first error was considered in this and all subsequent analyses. Interrater reliability was very high; the 59 computed unweighted kappa coefficients ranged from .78 to 1.00 (Mdn = 1.00). The weighted kappas ranged from .82 to 1.00 (Mdn = 1.00).

As expected, the total score on the Boston Naming Test (number correct out of 60 items) was significantly higher for the healthy controls (M = 55.5, SD = 3.6, range = 46–60) than for the very mildly demented subjects (M = 41.12, SD = 12.74, range = 8–60), whose total score in turn, was higher than that of the mildly demented individuals (M = 29.58, SD = 14.5, range = 6–56), $F(2, 71) = 35.52$, $p < .0001$. Overall, the control group made a total of 113 errors, the very mildly demented group made 456 errors, and the mildly demented group made 728 errors.

The mean ratings and standard deviations of the 17 error types made by the three groups are displayed in Table 3. Also included in Table 3 are the means and standard deviations for the no-content, linguistic, and perceptual categories, collapsed across error type within each category. All error types produced main effects of group (all ps < .05, one-way analyses of variance [ANOVAs]) with the exception of Types 3 and 10 within the linguistic category and Types 14, 16, and 17 within the perceptual category.

To determine whether the category of error type differentially changed across CDR level, we conducted a 3 (CDR = 0, .5, 1) x 3 (no-content, linguistic, perceptual) mixed-factor ANOVA. This analysis yielded a significant main effect of group, $F(2, 71) = 35.52$, $p < .0001$,

[4]We are not arguing that there is an exclusive correspondence between perceptual, linguistic, and no-content errors and perceptual, lexical, and semantic mechanisms. Clearly, an error in one domain may also reflect the contribution of difficulties in other domains. We are simply arguing that the priority of these underlying mechanisms with respect to these error types do produce the above-mentioned correspondence. As seen later in the Results section, subsequent analyses converge on this correspondence.

Table 3
Mean Ratings and Standard Deviations of Errors on the Boston Naming Test as a Function of Clinical Dementia Rating (CDR)

Type of error	Normal (CDR = 0) *M*	Normal (CDR = 0) *SD*	Very mildly demented (CDR = 0.5) *M*	Very mildly demented (CDR = 0.5) *SD*	Mildly demented (CDR = 1.0) *M*	Mildly demented (CDR = 1.0) *SD*
No content						
Empty phrase	0.44	0.65	2.24	2.85	7.54	8.31
No interpretation possible	0.32	0.69	0.92	1.38	3.04	3.93
Total	0.76	0.97	3.16	3.30	10.58	9.47
Linguistically related						
Phonological	0.16	0.37	0.56	0.65	0.67	1.12
Semantic						
Same category	1.12	1.39	3.76	2.37	3.04	1.97
Super- or subordinate category	0.20	0.41	1.44	2.18	1.75	2.09
Function	0.64	0.70	2.00	1.71	2.42	2.21
Attribute	0.08	0.28	0.72	1.02	2.08	1.86
Context	0.88	1.09	2.44	2.24	3.58	3.22
Description	0.12	0.33	0.88	1.17	0.67	0.87
Acoustic	0.00	0.00	0.00	0.00	0.00	0.00
Pantomime or gesture	0.08	0.28	0.36	0.76	1.38	1.56
Total	3.28	2.44	12.16	6.55	15.58	7.20
Perceptually related						
Whole	0.24	0.44	1.88	2.83	1.96	2.71
Part	0.24	0.60	0.68	0.95	1.38	1.31
Perspective	0.00	0.00	0.12	0.44	0.13	0.34
Function	0.00	0.00	0.04	0.20	0.33	0.56
Attribute	0.00	0.00	0.08	0.28	0.17	0.48
Context	0.00	0.00	0.12	0.33	0.21	0.59
Total	0.48	0.92	2.92	4.14	4.17	3.91

Note. The range of the no-content, linguistic, and perceptual errors (totals) for the normal control group were 0–3, 0–10, and 0–3, respectively. For the very mildly demented group, the ranges were 0–10, 0–24, and 0–15, respectively, and for the mildly demented group, the ranges were 0–34, 3–33, and 0–14, respectively.

indicating that errors increased with the severity of dementia, and a significant main effect of error category, $F(2, 142) = 61.08, p < .0001$, indicating that linguistic errors were more common than no-content errors, which in turn were more common than perceptual errors. More important for the present discussion, this analysis also yielded a highly significant

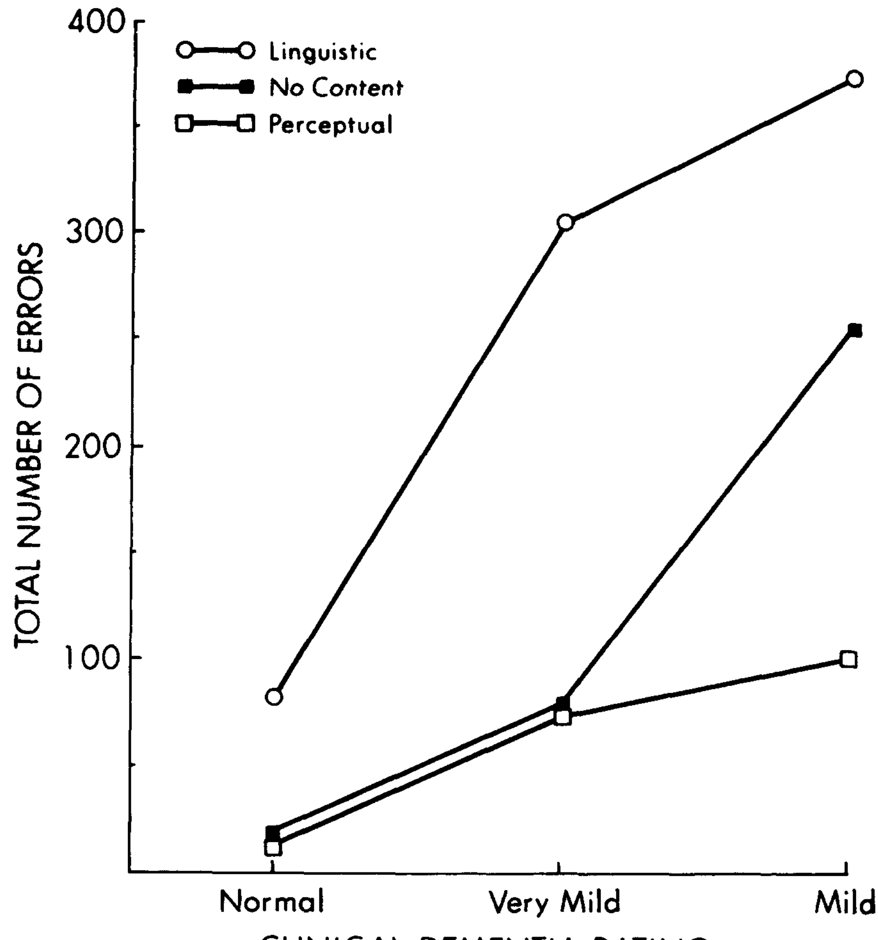

Figure 1. Total number of errors on the Boston Naming Test as a function of type of error and severity of dementia.

interaction between group and error category, $F(4, 142) = 9.12$, $p < .0001$. This interaction is graphically presented in Figure 1. Planned comparisons indicated that there was a reliable increase in perceptual errors between the control and the very mildly demented individuals ($p < .05$), whereas the difference between the very mildly demented and mildly demented individuals did not reach significance. On the other hand, for the linguistic errors there was a very large increase between the controls and the very mildly demented individuals ($p < .0001$), along with a smaller, but reliable, increase between the very mildly demented and the mildly demented individuals ($p < .01$). Finally, for the no-content errors, there was a large increase in errors between the very mildly and mildly demented individuals ($p < .0001$), whereas the difference between the control subjects and the very mildly demented individuals did not reach significance. Thus, the pattern of error types did change across severity of dementia, with linguistic errors occurring relatively early in the disease, and no-content errors appearing relatively late in the disease. The perceptual errors were relatively rare and, if anything, appeared to occur relatively early in the disease.

The data in Figure 1 support the notion that there was a differential change in the type of errors across the different levels of dementia. Assuming that linguistic errors primarily reflect a breakdown in lexical information, whereas no-content errors reflect some involvement of the semantic system, the pattern displayed in Figure 1 suggests that, early in SDAT, there is still access to semantic representations, with some breakdown in lexical access processes. At the mild stage of dementia, however, there appears to be an increasing involvement of the semantic system, as reflected by the increase in no-content errors.

Four-Alternative Multiple-Choice Test

To further explore the nature of the different error types across dementia level, we now present analyses from the four-alternative multiple-choice test. If, indeed, semantic structures are intact and the mildly demented individuals primarily had difficulty with lexical retrieval, then one would expect these individuals to benefit considerably from the presentation of the correct lexical item on a recognition test. To ensure that the particular items that produced the errors across CDR levels did not change in difficulty, we analyzed only the set of pictures that produced at least one error for each CDR level. This resulted in a set of 34 pictures. Table 4 presents the percentage correct on the four-alternative multiple-choice test as a function of CDR and error type.

First, as shown in Table 4, all of the healthy control subjects correctly recognized all of the words that they were unable to retrieve in response to the target pictures. Thus, it appears that a difficulty in lexical processing accounts for the errors produced by the healthy control subjects, because these individuals could always correctly select the correct alternative on the multiple choice test. Second, accuracy was also quite high for the very mildly demented individuals ($M = 91\%$). Thus, the errors produced by the very mildly demented individuals probably also involved a breakdown in lexical processing. A different pattern was produced by the mildly demented individuals, whose overall recognition accuracy was only 64%. Moreover, recognition accuracy was especially low for the no-content errors (55%). Thus, the mildly demented individuals produced a rather large impairment in performance, even when they were presented with the correct alternative on a four-alternative multiple-choice test. This pattern supports the notion that simple lexical retrieval is not the sole problem for mildly demented individuals.

Table 4

Recognition Performance (Mean Percentage Correct) as a Function of Clinical Dementia Rating (CDR) and Type of Error

Type of error	Normal (CDR = 0)	Very mildly demented (CDR = 0.5)	Mildly demented (CDR = 1.0)
No content	100	91	55
Linguistically related	100	93	76
Perceptually related	100	89	63

Note. For the normal group, $n = 113$ observations; for the very mildly demented group, $n = 378$ observations; and for the mildly demented group, $n = 532$ observations.

Because the control subjects were obviously performing at ceiling levels, we included only the very mildly and mildly demented individuals in a 3 (error type) x 2 (CDR) ANOVA on recognition performance (percentage correct). This ANOVA yielded a main effect of CDR, $F(1, 170) = 61.97, p < .0001$, error type, $F(2, 170) = 5.04, p < .01$, and a CDR x Error Type interaction, $F(2, 170) = 2.96, p < .05$. Planned comparisons indicated that the very mildly demented individuals were reliably more accurate on each of the error types than were the mildly demented individuals (all $ps < .005$). Moreover, for the mildly demented individuals, recognition accuracy was reliably lower for the no-content errors than for the linguistic errors ($p < .0001$); however, the difference between the perceptual errors and the no-content errors did not reach significance ($p < .28$). The rather low recognition performance on the no-content errors for the mildly demented individuals provides further support for the argument that the large increase in no-content errors across the very mildly demented and the mildly demented individuals (see Figure 1) does not simply involve a breakdown in lexical processing but rather is more likely to include semantic structures.

Semantic Rating Analysis

The analyses described in the previous section suggest that there is an increasing involvement of semantic structures in mildly demented individuals. This argument is supported both by the error-type analyses in the Boston Naming Test and the results from the four-alternative multiple-choice test. To provide further information regarding this breakdown, we had six naive individuals rate the linguistic errors on a scale that reflected the adequacy of the response in capturing the meaning of the picture. These six raters were told that the 60 line drawings were given to a group of older adults and that the incorrect responses were recorded when subjects could not correctly name the picture. The raters were asked to evaluate how fully the erroneous responses captured the meaning of the line drawing. The raters were given a 5-point scale on which 1 indicated *best—a dictionary definition*, 2 indicated *moderately good—good understanding but not best*, 3 indicated *adequate—clearly aware of the general idea*, 4 indicated *moderately poor—poor understanding but some relationship*, and 5 indicated *worst—no relationship*. The raters were encouraged to use all levels of the rating scale.

The raters were then given the randomly ordered pictures along with the randomly ordered linguistic errors produced across CDR levels for each picture. A Cronbach's alpha of .87 indicated that the raters understood the task and were consistent in their ratings of the errors.

To ensure that any differences across the three dementia levels were not simply due to differences in the particular pictures that produced errors, we included only the 28 pictures that produced at least one linguistic error across all three dementia levels. The results of this rating procedure were quite clear: The mean ratings for the control subjects (3.1) did not differ from those for the very mildly demented individuals (3.1), but both of these mean ratings were different from the mean ratings of the mildly demented individuals (3.5). The main effect of CDR level was reliable, $F(2, 81) = 4.20, p < .02$. Hence, these data indicate that the linguistic errors produced by the mildly demented individuals were indeed rated lower on the extent to which the errors captured the meaning of the picture than were the linguistic errors of both the control and very mildly demented individuals. This again supports the notion that there is an increasing involvement of semantic structures in mildly demented individuals and that there is relatively little difference along this dimension between the control and the very mildly demented individuals.

Correlations Between Boston Naming Test Performance and Psychometric Measures

The mean scores and standard deviations for the various psychometric tests are displayed in Table 5. As expected, each of these tests provided significant main effects of CDR level. (For more detailed discussion of the degree to which each of these tests discriminates between different stages of the disease, see Storandt & Hill, 1989). The important point to note for the present discussion is that the mildly demented individuals appear to have performed at floor on the Logical Memory test, whereas the healthy control individuals appear to have performed at ceiling level on the Benton Copy Test. Hence, the importance of changes in correlations across dementia levels is limited by the problem of sufficient ranges in performance on these two tasks. The remaining tasks, however, do appear to have similar ranges in performance and hence may provide useful information about correlations between changes in psychometric performance with Boston Naming Test performance.

Table 6 displays the correlations between each of these psychometric tests and overall Boston Naming Test performance. There are three points to note. First, for the healthy control group, the only task that significantly correlated with Boston Naming Test performance was the WAIS Information subtest. This should be expected because the WAIS Information test taps preexisting verbal information. There was also some correlation between Boston Naming Test performance and performance on the WMS Mental Control subtest and the Word Fluency test (both $ps < .10$). The correlation between Word Fluency and Boston Naming Test performance is especially noteworthy because it supports the notion that difficulty in lexical retrieval processes underlies the healthy controls' errors in the Boston

Table 5
Mean Scores and Standard Deviations on the Psychometric Tests as a Function of Clinical Dementia Rating (CDR)

	Normal (CDR = 0)		Very mildly demented (CDR = 0.5)		Mildly demented (CDR = 1.0)	
Test	*M*	*SD*	*M*	*SD*	*M*	*SD*
Wechsler Memory Scale						
Associate Memory	12.7	3.5	8.2	3.2	5.7	1.9
Logical Memory	9.3	2.8	4.7	2.3	1.2	1.3
Digit Span (forward and backward)	12.4	2.1	10.1	2.7	8.5	2.2
Mental Control	7.2	2.1	5.6	2.3	4.6	2.5
Benton Visual Retention Test	7.0	3.2	12.1	4.3	17.7	5.5
Benton Copy Test	0.6	1.0	1.6	2.8	3.8	3.6
Wechsler Adult Intelligence Scale						
Information	21.5	3.6	13.4	6.2	7.4	4.0
Comprehension	22.4	3.4	17.4	5.6	11.0	5.1
Block Design	30.4	8.1	18.1	11.4	11.3	11.4
Digit Symbol	45.4	9.6	31.0	9.4	20.2	12.6
Trail Making A	41.6	12.2	68.2	30.6	111.0	53.2
Word Fluency	29.3	10.7	18.8	9.4	12.8	8.8

Naming Test. It is also intriguing that none of the remaining psychometric tests was significantly correlated with Boston Naming Test performance in the healthy older adults. This is consistent with a recent article by Albert et al. (1988), who found that basic episodic memory performance did not significantly correlate with Boston Naming Test performance. Of course, one must be relatively cautious in interpreting the lack of significant correlations in the control group because, as noted previously, these individuals produced a rather narrow range of scores in Boston Naming Test performance (i.e., M = 55.48, SD = 3.56).

The second observation to make from Table 6 is that there were rather strong correlations between Boston Naming Test performance and a wide range of the psychometric tests for the very mildly demented individuals. Of particular interest are the strong correlations between basic intelligence (as measured by the WAIS subtests), memory performance (as measured by the WMS Associate Learning and Logical Memory subtests and the Benton Visual Retention Test), and perceptual motor performance (as measured by the Trail Making A), along with the basic breakdown in performance on the WMS Mental Control subtest and the Word Fluency test. These results suggest that poor Boston Naming Test performance is only one of a myriad of symptoms of a more generalized cognitive deficit in very mild dementia. This is noteworthy because at least early on in the disease it appears that there is a general breakdown in cognitive performance rather than a limited breakdown in specific cognitive mechanisms, at least as these mechanisms are related to Boston Naming Test performance.

The third point to note in Table 6 is that intellectual performance as measured by the WAIS Information and Comprehension subtests was highly correlated with Boston Naming Test performance for the mildly demented individuals. These correlations are consistent with the notion that there is a breakdown in basic knowledge representations (semantic memory) at this level of the disease. The only remaining reliable correlation for the mildly demented

Table 6
Correlations Between Performance on Psychometric Tests and the Boston Naming Test (Total Number of Errors) as a Function of Clinical Dementia Rating (CDR)

Test	Normal (CDR = 0)	Very mildly demented (CDR = 0.5)	Mildly demented (CDR = 1.0)
Wechsler Memory Scale			
Associate Memory	–.06	–.52**	–.36
Logical Memory	–.22	–.40*	–.16
Digit Span (forward and backward)	–.08	–.02	–.23
Mental Control	–.37	–.57**	–.42*
Benton Visual Retention Test	.12	.50*	–.05
Benton Copy Test	.28	.19	–.09
Wechsler Adult Intelligence Scale			
Information	–.40*	–.69***	–.58**
Comprehension	–.05	–.63**	–.57**
Block Design	–.22	–.56**	.22
Digit Symbol	–.03	–.47*	.07
Trail Making A	.21	.48*	.04
Word Fluency	–.34	–.44*	–.22

* $p < .05$. ** $p < .01$. *** $p < .0001$.

individuals was between Boston Naming Test performance and performance on the WMS Mental Control subtest.

Discussion

The results of the present study are quite clear. First, the present results are consistent with recent studies indicating that the primary breakdown in SDAT individuals, at least with respect to Boston Naming Test performance, is in semantic and lexical processing. Perceptual errors were relatively rare and, in some cases, probably reflected the involvement of a semantic or lexical system. For example, one individual gave the description *fellow who goes underground* in response to the picture of a beaver. It is quite possible that this error was due to the access of the rodent category, via beaver, and the retrieval of information available in that category (i.e., many rodents go underground). The important point for the present discussion is that, even if all of these errors had a true perceptual basis, they were relatively rare and did not disproportionately increase in likelihood as a function of the severity of the dementia, even though overall error rates on the Boston Naming Test increased dramatically. Thus, our results are consistent with the growing evidence that errors in confronation naming are not primarily due to a perceptual breakdown (e.g., Bayles & Tomoeda, 1983; Huff et al., 1986; Martin & Fedio, 1983; Smith et al., 1989; Wilson, Kazniak, Fox, Garron, & Ratusnik, 1981).

The second major finding in this study is that there was a differential progression of no-content and linguistic errors across the three groups of subjects. First, consider the linguistic errors. There was a large increase in linguistically related errors between the control and the very mildly demented individuals, with a relatively smaller increase between the very mildly demented and mildly demented individuals. Although these errors may reflect some semantic breakdown, the fact that the linguistic error responses were often related to the target suggests that at least part of this effect was due to a failure to retrieve the correct lexical representation. On the other hand, the no-content errors did not dramatically increase until subjects reached the mild level of dementia. It appears that mildly demented individuals are more likely to lose access to both the lexical and core semantic information that are necessary to give even a lexically related response. This pattern clearly suggests an increasing involvement of semantic structures as Alzheimer's disease progresses.

The general distinction between lexical and semantic breakdowns across severity of dementia was also supported by two further analyses. First, the four-alternative multiple-choice data indicated that the normal and very mildly demented individuals could correctly recognize the correct alternative. For the mildly demented individuals, however, performance was relatively poor even when they were presented with the correct lexical item on a multiple choice test. This was particularly salient for the no-content errors. Thus, the rather large increase in no-content errors in the mildly demented individuals, and their poor performance on the recognition test for these same items, clearly supports the notion that these errors involve more than simply an inability to retrieve lexical information.

The second piece of evidence supporting a semantic loss were the ratings of the errors produced by the different groups. There was no difference between the control and very mildly demented individuals. However, the mildly demented individuals produced errors that did not as adequately capture the meaning of the pictures as did the errors of the other two groups.

These arguments were also supported by the correlational analyses. For the healthy older adults, there was a relationship between performance on the Boston Naming Test and performance on the WAIS Information subtest, along with some evidence of a relationship between Boston Naming Test performance and performance on the Word Fluency subtest of the WMS. This pattern could be viewed as being consistent with arguments made in the literature about healthy older individuals' experiencing a breakdown in lexical retrieval processes (e.g., Bowles & Poon, 1985).

Among the very mildly demented individuals, there were significant correlations between the measures of lexical processing and Boston Naming Test performance, with the inclusion of a host of other variables. For the very mildly demented individuals, Boston Naming Test performance reflected only one of a myriad of cognitive deficits, including poor episodic memory (as assessed by the WMS Associate Memory subtest and the Benton Visual Retention Test). This is intriguing because it suggests that there is a widespread breakdown in performance in the very mildly demented individuals, as opposed to a rather limited breakdown of a particular underlying structure.

Finally, for the mildly demented individuals, there was a significant relationship between Boston Naming Test performance and performance on the WAIS Information and Comprehension subtests, along with a marginally reliable relationship between Boston Naming Test performance and performance on the WMS Mental Control subtest.

Overall, the present results suggest that the deficit in confrontation naming performance in SDAT reflects a progressive involvement of semantic structures. These results are consistent with the conceptual frameworks described by Flicker, Ferris, Crook, and Bartus (1987) and Warrington (1975). According to these researchers, specific attributes or subordinate information is especially sensitive to SDAT, whereas superordinate information stays relatively intact. In fact, Flicker et al. suggested that in the early stages of SDAT there appears to be restricted access to the specific attributes within a category. This is quite consistent with the increased involvement of the lexical system in the very mildly demented individuals. As the disease progresses, however, there appears to be a greater involvement of more central categorical information, yielding (a) an increase in no-content errors, (b) rather poor multiple-choice recognition performance, and (c) the production of error responses that were rated as relatively poor in capturing the meaning of the pictures.

In this light, the present results are most consistent with a framework in which the progression of SDAT involves a decreasing degree of specificity of language structures and processes. Early in the disease, there is a loss of the specific verbal label for a concept, followed by what seems to be a breakdown in local semantic attributes, until eventually there is some involvement of core semantic structures and processes. In fact, Balota and Duchek (1991) recently reported that lexical context (e.g., music) is less likely to disambiguate homographs (e.g., *organ* referring to a musical instrument vs. *organ* referring to a bodily organ) in mildly demented individuals. The specification of this language deterioration should lead to a a better understanding of the language deficits exhibited by these individuals and, ultimately, better diagnostic criteria for the early stages of SDAT. For example, the clinician can be more certain of a diagnosis of mild SDAT if poor performance on the Boston Naming Test (in the context of poor episodic memory performance) is the result of a loss of semantic information as revealed by numerous *I don't know* answers.

Acknowledgments

Preparation of this article was supported in part by the Norman J. Stupp Foundation and by National Institute on Aging Grants AGO5681, AGO3911, and RO15427A.

A version of this research was presented at the 41st Annual Meeting of the Gerontological

Society of America, San Francisco, CA.

We thank Richard Ferraro, John Morris, and especially Dorothy Edwards for their help on various aspects of this project.

References

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

Appell, J., Kertesz, A., & Fisman, M. (1982). A study of language functioning in Alzheimer patients. *Brain and Language, 17,* 73–91.

Armitage, S. G. (1946). An analysis of certain psychological tests used in the evaluation of brain injury. *Psychological Monographs, 60* (1, Whole No. 277), 1–48.

Balota, D. A., & Chumbley, J. I. (1984). Are lexical decisions a good measure of lexical access? The role of the neglected decision stage. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 340–357.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Access and/or production? *Journal of Memory and Language, 24,* 89–106.

Balota, D. A., & Duchek, J. M. (1988). Age related differences in lexical access, spreading activation, and simple pronunciation. *Psychology and Aging, 3,* 84–93.

Balota, D. A., & Duchek, J. M. (1989). Spreading activation in episodic memory: Further evidence for age independence. *Quarterly Journal of Experimental Psychology, 41*A, 849–876.

Balota, D. A., & Duchek, J. M. (1991). Semantic priming effects, lexical repetition effects, and contextual disambiguation effects in healthy aged individuals and individuals with senile dementia of the Alzheimer type. *Brain and Language, 40,* 181–201.

Barker, M. G., & Lawson, J. S. (1968). Nominal aphasia in dementia. *British Journal of Psychiatry, 114,* 1351–1356.

Bayles, K. A. , & Tomoeda, C. K. (1983). Confrontation naming impairment in dementia. *Brain and Language, 19,* 98–112.

Becker, J. T., Huff, J., Nebes, R. D., Holland, A., & Boller, F. (1988). Neuropsychological function in Alzheimer's disease. *Archives of Neurology, 45,* 263–268.

Benton, A. L. (1963). *The Revised Visual Retention Test: Clinical and experimental applications.* New York: Psychological Corp.

Berg, L. (1984). Clinical dementia rating. *British Journal of Psychiatry, 145,* 339.

Berg, L. (1988). Clinical dementia rating (CDR). *Psychopharmacology Bulletin, 24,* 637–639.

Berg, L. (1988). Mild senile dementia of the Alzheimer type: Diagnostic criteria and natural history. *Mount Sinai Journal of Medicine, 55,* 87–96.

Berg, L., Hughes, C. P., Coben, L. A., Danziger, W. L., Martin, R. L. & Knesevich, J. (1982). Mild senile dementia of Alzheimer type: Research diagnostic criteria, recruitment, and description of a study population. *Journal of Neurology, Neurosurgury, and Psychiatry, 45,* 962–968.

Botwinick, J., Storandt, M., & Berg, L. (1986). A longitudinal behavioral study of senile dementia of the Alzheimer type. *Archives of Neurology, 43,* 1124–1127.

Bowles, N. L., & Poon, L. W. (1985). Aging and retrieval of words in semantic memory. *Journal of Gerontology, 40,* 71–77.

Burke, W. J., Miller, J. P., Rubin, E. H., Morris, J. C., Coben, L. A., Duchek, J., Wittels, I. G., & Berg, L. (1988) Reliability of the Washington University Clinical Dementia Rating. *Archives of Neurology, 45,* 31–32.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American Heritage word frequency book.* New York: American Heritage Publishing Co. & Houghton Mifflin.

Faber-Langendoen, K., Morris, J. C., Knesevich, J. W., LaBarge, E., & Berg, L. (1988). Aphasia in senile dementia of the Alzheimer type. *Annals of Neurology, 23,* 365–370.

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

Forster, K. L. (1978). Accessing the mental lexicon. In E. Walker (Ed.), *Explorations in the biology of language* (pp. 139–176). Montgomery, VT: Bradford Books.
Goodglass, H., & Kaplan, E. (1983). *Boston Naming Test*. Philadelphia, PA: Lea & Febiger.
Harrold, R. M. (1988, August). *Object naming in Alzheimer's disease: What is the cognitive deficit?* Paper presented at the 96th Annual Meeting of the American Psychological Association, Atlanta, GA.
Hollingshead, A. B. (1957). *Two Factor Index of Social Position*. New Haven, CT: Yale University.
Huff, F. J., Corkin, S., & Growdon, J. H. (1986). Semantic impairment and anomia in Alzheimer's disease. *Brain and Language*, *28*, 235–249.
Hughes, C. P., Berg, L., Danziger, W. L., Coben, L. A., & Martin, R. L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry*, *140*, 566–572.
Joynt, R. J. (1984). The language of dementia. In F. C Rose (Ed.), *Advances in neurology* (pp. 65–69). New York: Raven Press.
Kirshner, H. S., Webb, W. G., & Kelly, M. P. (1984). The naming disorder of dementia. *Neuropsychologia*, *22*, 23–30.
Knesevich, J., LaBarge, E., & Edwards, D. (1986). Predictive value of the Boston Naming Test in mild senile dementia of the Alzheimer type. *Psychiatry Research*, *19*, 155–161.
Knesevich, J., Toro, F., Morris, J., & LaBarge, E. (1985). Aphasia, family history, and the longitudinal course of senile dementia of the Alzheimer type. *Psychiatry Research*, *14*, 255–263.
Lawson, J. S., & Barker, M. G. (1968). The assessment of nominal dysphasia in dementia: The use of reaction time measures. *British Journal of Medical Psychology*, *41*, 414–441.
Martin, A., & Fedio, P. (1983). Word production and comprehension in Alzheimer's disease: The breakdown of semantic knowledge. *Brain and Language*, *19*, 124–141.
Mayeux, R., Stern, Y., & Spanton, S. (1985). Heterogeneity in dementia of the Alzheimer type: Evidence of subgroups. *Neurology*, *35*, 453.
Morris, J. C., McKeel, D. W., Fulling, K., Torack, R., & Berg, L. (1988). Validation of clinical diagnostic criteria for Alzheimer's disease. *Annals of Neurology*, *24*, 17–22.
Morris, J. C., McKeel, D. W., Jr., Price, J. L., LaBarge, E., Vendegna, T. R., Rubin, E. H., Berg, L., & Ball M. J. (1988). Very mild senile dementia of the Alzheimer type (SDAT). *Neurology*, *38*(Suppl. 1), 227.
Nebes, R. (1989). Semantic memory in Alzheimer's disease. *Psychological Bulletin*, *106*, 377–394.
Rochford, G. (1971). A study of naming errors in dysphasic and demented patients. *Neuropsychologia*, *9*, 437–443.
Rubin. E. H., Morris, J. C., Grant, E. A., & Vendegna, T. (1989). Very mild senile dementia of the Alzheimer type: I. Clinical assessment. *Archives of Neurology*, *46*, 379–382.
Skelton-Robinson, M., & Jones, S. (1984). Nominal dysphasia and the severity of senile dementia. *British Journal of Psychiatry*, *145*, 168–171.
Smith, S., Murdoch, B., & Chenery, H. (1989). Semantic abilities in dementia of the Alzheimer type. *Brain and Language*, *36*, 314–324.
Snodgrass, J. G. (1984). Concepts and their surface representations. *Journal of Verbal Learning and Verbal Behavior*, *23*, 3–22.
Storandt, M., & Hill, R. D. (1989). Very mild senile dementia of the Alzheimer type: II. Psychometric test performance. *Archives of Neurology*, *46*, 383–386.
Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Columbia University Press.
Thurstone, L. L., & Thurstone, L. G. (1949). *Examiner manual for the SRA Primary Mental Abilities Test*. Chicago: Science Research Associates.
Warren, L. R., & Morton, J. (1982). The effects of priming on picture recognition. *British Journal of Psychology*, *57*, 147–154.
Warrington, E. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, *27*, 635–657.
Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. New York: Psychological Corp.

Wechsler, D., & Stone, C.P. (1973). *Manual for the Wechsler Memory Scale*. New York: Psychological Corp.

Wilson, R. S., Kaszniak, A. W., Fox, J. H., Garron, D. C., & Ratusnik, D. L. (1981, February). *Language deterioration in dementia*. Paper presented at the Ninth Annual Meeting of the International Neuropsychological Society, Atlanta, GA.