

Semantic Memory in Alzheimer's Disease and the Frontotemporal Dementias: A Longitudinal Study of 236 Patients

Timothy T. Rogers

Medical Research Council (MRC) and University
of Wisconsin-Madison

Adrian Ivanoiu

Catholic University of Louvain

Karalyn Patterson

Medical Research Council (MRC)

John R. Hodges

Medical Research Council (MRC) and University of Cambridge

Using semantic dementia (SD) as a reference point, the authors assessed semantic memory in four other neurodegenerative disorders: progressive nonfluent aphasia (PNFA), frontal variant frontotemporal dementia (fvFTD), Alzheimer's disease (AD), and posterior cortical atrophy (PCA). Individuals with SD were more impaired than other groups on semantic measures and showed a characteristic pattern across tasks: category fluency (CF) worse than letter fluency (LF), naming worse than comprehension, and visual and verbal comprehension equally affected, suggesting disruption to an amodal semantic system. Individuals with AD demonstrated a similar pattern to a milder degree. Although PNFA, fvFTD, and PCA groups had abnormal scores (relative to controls) on most semantic measures, their differing patterns across measures indicate that the apparent semantic impairment in these conditions is largely secondary to other factors.

Keywords: semantic memory, frontotemporal dementia, comprehension impairment, visual processing impairments, cognitive assessment measures

Semantic memory is the form of memory that allows humans to understand the meanings of words, objects, and events; to name and to produce meaningful statements; and to behave in a manner that reflects knowledge about the “kinds of things” that exist in the environment. This study assesses the status of semantic memory in five different neurodegenerative syndromes, using a battery of common semantic tests: verbal fluency, naming, word-picture matching (WPM), and the picture and word versions of the Pyramids and Palm Trees test of semantic association. All of these measures are known to be sensitive to semantic impairment, but it is also clear that performance on each may suffer as a consequence of deficits to other nonsemantic cognitive faculties. For example, disruption to the mechanisms of speech production may yield deficits in fluency and naming with or without significant coinci-

dent semantic impairment, difficulties in executive function may impair verbal fluency or WPM tasks with multiple response alternatives, and visual impairment may give rise to differentially poor performance on the picture version of the Pyramids and Palm Trees test. The assessment of semantic abilities in neurodegenerative syndromes thus presents two related challenges.

First, abnormal performance on individual semantic tests (or even on a battery of such tests) may reflect impairment to the semantic memory system proper; or it may reflect the abnormal functioning of other cognitive faculties that provide input to or encode output from the semantic system. For instance, individuals with AD often perform poorly on semantic tasks such as CF and WPM (Chertkow & Bub, 1990; Hodges & Patterson, 1995; Salmon, Heindel, & Lange, 1999); however, they also show deficits in assessments of episodic memory (Welsh, Butters, Hughes, Mohs, & Heyman, 1992), executive function (Perry & Hodges, 2000), and in some cases, visual perception (Caine & Hodges, 2001). In the context of these deficits, it is not clear whether impairment to WPM, verbal fluency, naming, and so forth, really indicates dysfunction in the semantic system (Bayles, Tomoeda, Kaszniak, & Trosset, 1991; Kempler, Anderson, & Henderson, 1995; Martin, Brouwers, Cox, & Fedio, 1985).

Second, in cases where the clearest evidence exists of semantic memory impairment, it may be difficult to determine whether these coexist with other, nonsemantic cognitive deficits. For example, significant impairments on a broad range of semantic tests, alongside relative sparing of other cognitive function, characterizes SD (Hodges, Patterson, Oxbury, & Funnell, 1992; Snowden, Goulding, & Neary, 1989). A reasonable working hypothesis is that this pattern of performance reflects the degradation of the semantic memory system; however, some individuals with SD show a strikingly severe degree of anomia, seemingly out of proportion

Timothy T. Rogers, MRC Cognition and Brain Sciences Unit, Cambridge, England; Department of Psychology, University of Wisconsin-Madison, Madison, WI; Adrian Ivanoiu, Department of Neurology, Memory Clinic and Neuropsychological Rehabilitation Centre, Catholic University of Louvain, Brussels, Belgium; Karalyn Patterson and John R. Hodges, MRC Cognition and Brain Sciences Unit, Cambridge, England.

The first two authors (Timothy T. Rogers and Adrian Ivanoiu) contributed equally to the work described in this paper. The order listed for these authors is arbitrary. We would like to extend our thanks to Ian Nimmo-Smith at the MRC Cognition and Brain Sciences Unit for helpful advice on the application of mixed linear models to our data. Both Timothy T. Rogers and Karalyn Patterson were supported in part by an NIMH Interdisciplinary Behavioural Science Center grant.

Correspondence concerning this article should be addressed to Timothy T. Rogers, Department of Psychology, University of Wisconsin-Madison, Madison WI. E-mail: ttrogers@wisc.edu; and Dr Adrian Ivanoiu, MD, Department of Neurology, Saint Luc University Hospital, 10 Avenue Hippocrate, B-1200, Brussels, Belgium. E-mail: ivanoiu@nefy.ucl.ac.be

with the magnitude of their semantic impairment (e.g., Lambon Ralph, McClelland, Patterson, Galton, & Hodges, 2001). Because anomia is characteristic of semantic impairment, it is difficult to determine whether these serious naming difficulties arise solely from the degradation of the semantic system, or whether they also reflect the concurrent deterioration of a lexical or phonological output system. More generally, in individuals with clear semantic impairments, it may be difficult to determine whether nonsemantic mechanisms of visual and verbal reception and expression are normal.

These challenges have led to considerable dispute regarding the neural and cognitive underpinnings of human semantic memory. Although some researchers believe that the neuropathology of AD produces serious semantic memory deficits (e.g., Chan, Salmon, Butters, & Johnson, 1995; Chertkow & Bub, 1990; Hodges & Patterson, 1995; Hodges, Salmon, & Butters, 1992; Martin, 1992), others have argued that these apparent semantic impairments actually arise from difficulties of access to or retrieval from a relatively intact semantic store (see Bayles et al., 1991; Glosser & Friedman, 1991; Nebes, 1989; Nebes & Brady, 1990; Ober, Shenaut, & Reed, 1995). The position one takes on this issue obviously has important implications for a neuropsychological theory of semantic memory: If the semantic store is intact in AD, then neural systems that are structurally and metabolically intact in the disorder presumably support it; if it is degraded, then the reverse must be true. Similarly, we (and others) have suggested that SD reflects the deterioration of a central, amodal semantic system (Lambon Ralph et al., 2001; Rogers et al., 2004) located in the anterior and inferior temporal lobe that is the consistent locus of atrophy in the disorder (Mummery et al., 2000). Other researchers, however, have argued that the condition actually reflects the degeneration of relatively independent verbal and visual knowledge systems in the temporal lobes (e.g., Mesulam, Grossman, Hillis, Kertesz, & Weintraub, 2003). Here again, one's theory about the neural underpinnings of semantic memory depends critically on how one interprets the comparative performance of individuals on different semantic tasks, each of which might be influenced by other nonsemantic impairments.

Thus to assess semantic capabilities in neurodegenerative disease, it is important to disentangle the different factors that may contribute to impaired performance on any given semantic measure. One way to accomplish this end is to compare performance on tests that are typically considered to be *semantic* or *nonsemantic* across neurodegenerative conditions that affect different aspects of cognition. However, no large-scale comparison exists between groups of individuals with different cortical neurodegenerative syndromes on tests of semantic memory and other standard neuropsychological assessments. The current paper describes such a comparison among five groups of individuals in an effort to assess the integrity of semantic memory in each.

Three of the patient groups had variants of frontotemporal lobar degeneration (FTD) (Grossman, 2002; Hodges & Miller, 2001; Neary et al., 1998). Individuals with the temporal variant of FTD, usually referred to as SD (Hodges, Graham, & Patterson, 1995; Snowden et al., 1989), have degraded knowledge about the meanings of words and objects, whereas executive functions, visuospatial and perceptual abilities, and at least some aspects of episodic memory are all relatively well preserved. Brain imaging reveals a characteristic pattern of atrophy involving the anterior portion of the temporal lobes, almost always asymmetrical and frequently

with predominance on the left side (Chan et al., 2001; Galton et al., 2001b; Mummery et al., 2000), but inevitably becoming bilateral as the disease progresses.

The second variant, PNFA, produces a severe disruption of speech output characterized by phonological and syntactic errors (Grossman, 2002; Mesulam, 1982). In such cases it is possible to observe mild deficits in semantic tasks, although these are never as pronounced as in SD (Croot, Patterson, & Hodges, 1998). A recent FDG-PET study indicates that the major area of dysfunction is centered in the left inferior frontal lobe, particularly the anterior insula (Nestor, Graham, et al., 2003).

The third major frontal/behavioral variant (fvFTD) produces changes in personality and social behavior, followed by executive dysfunction without major semantic impairment (Perry & Hodges, 2000; Rahman, Sahakian, Hodges, Rogers, & Robbins, 1999). Neuropsychological and structural imaging studies, together with postmortem pathology (Broe et al., 2003), reveal atrophy of the ventromedial frontal region in the earlier stages of the disease, often progressing backward to involve the anterior temporal lobes in the later stages (Rosen et al., 2001).

Typical and atypical variants of AD constitute the fourth and fifth progressive disorders included in our comparison. Individuals in the typical group most often have severe impairment of episodic memory in keeping with early involvement of medial temporal lobe regions (Braak & Braak, 1991) and of other components (e.g., posterior cingulate, retrosplenial cortex) of the Papez circuit (Nestor, Fryer, Smielewski, & Hodges, 2003). Deficits of semantic memory and visuospatial processing are well documented in AD (Caine & Hodges, 2001; Chertkow & Bub, 1990; Hodges & Patterson, 1995); executive dysfunction is also a consistent feature that may affect performance on semantic tasks (Perry & Hodges, 1999, 2000). Individuals in the atypical PCA group, in addition to memory and executive impairments, show disproportionate difficulty with tasks tapping visuooperceptual and spatial abilities (Benson, Davis, & Snyder, 1988; Galton, Patterson, Xuereb, & Hodges, 2000). FDG-PET imaging reveals marked hypometabolism of occipitoparietal regions, especially on the right (Nestor, Caine, Fryer, Clarke, & Hodges, 2003).

The individuals included in the study thus were selected to represent (a) a progressive semantic syndrome (SD); (b) a syndrome that progressively disrupts verbal production (PNFA) and may therefore influence performance on semantic tasks such as CF and naming; (c) a syndrome that progressively disrupts executive functioning (fvFTD) and may be expected to affect semantic tasks that tax executive resources, such as verbal fluency; (d) a syndrome that progressively disrupts visual perception (PCA) and may influence performance on visual semantic tasks; and (e) a progressive syndrome that affects multiple cognitive faculties (AD). The comparison of these groups on standard semantic and nonsemantic measures accomplishes three important aims. First, it allows for assessment of the status of semantic memory in the different groups by taking SD as the canonical example of a semantic syndrome and looking to see whether other groups show a similar pattern of spared and impaired performance across tasks. Second, it allows for assessment of the impact of nonsemantic factors on tests that ostensibly measure semantic memory to determine how best to diagnose semantic impairment in a neurodegenerative context. Finally, it allows for the investigation of the

factors that underlie performance across the full complement of tests in service of deriving a functional characterization of the five different syndromes.

Method

Participants

Patient groups. A total of 236 patient volunteers (42 SD, 23 PNFA, 40 fvFTD, 113 AD, and 18 PCA) contributed to the results. All cases were initially screened in the Memory Disorders Clinic at Addenbrooke's Hospital, Cambridge, and then were enrolled in MRC-funded longitudinal studies based in the Department of Neurology under the direction of consultant neurologist and author, John R. Hodges. Initial diagnoses were established following neurological and neuropsychological assessment using standard tests, with diagnosis arrived at by the consensus of clinicians in the memory clinic. Because recruitment through the clinic spanned the years 1991 to 2001, a variety of standard tests were used over this period; however, the experimental tests described below were not used for diagnostic purposes. For the three variants of FTD (SD, PNFA, and fvFTD), we applied the international consensus criteria (Neary et al., 1998) with the exception that the term *FTD* was applied as a general superordinate label, rather than frontotemporal lobar degeneration, in keeping with previous studies by our group (Bozeat, Gregory, Lambon Ralph, & Hodges, 2000; Galton et al., 2001a; Hodges et al., 1999; Ikeda, Brown, Holland, Fukuhara, & Hodges, 2002; Perry & Hodges, 2000; Rahman et al., 1999). All patients with SD, PNFA, and fvFTD underwent structural brain imaging (magnetic resonance imaging [MRI] in the vast majority of cases), and the scans for all cases of SD showed asymmetric anterior temporal lobe atrophy (see Thompson, Patterson, & Hodges, 2003). Most of these cases have been enrolled into a brain donation program and, to date, neuropathological confirmation has been possible in 30 patients with FTD syndromes, the majority of whom had non-Alzheimer's pathology (see Hodges, Davies, Xuereb, Kril, & Halliday, 2003; Hodges et al., 2004).

Patients with AD fulfilled the National Institute of Neurological and Communications Disorders (NINCDS) AD criteria (McKhann et al., 1984) and had predominant impairment of episodic memory. PCA patients had progressive visuospatial and perceptual deficits in the context of abnormalities in the parieto-occipital region on structural or functional brain imaging (or on both) (Benson et al., 1988; Mendez, Ghajarian, & Perryman, 2002; Nestor, Caine, et al., 2003). The demographic characteristics of the participants included in the study are displayed in Table 1, together with the number of test rounds administered to each patient. On average three to four longitudinal sets of observations were recorded for cases in

the SD, PNFA, AD, and PCA groups, and one to two observations for individuals in the fvFTD group, with gaps ranging from 6 to 12 months between testing rounds.

Controls. Data on a range of tasks were also collected from 106 healthy controls from the MRC Cognition and Brain Sciences Unit volunteer panel, with a mean age of 70.5 ($SD = 8.7$). Demographic details from the control group also appear in Table 1. The control participants were selected to match the largest patient group (AD) for age and education.

Neuropsychological Assessment

Assessment of semantic memory focused on five tests regularly included in the Cambridge Semantic Memory Battery (Bozeat et al., 2000; Garrard, Patterson, Watson, & Hodges, 1998; Garrard et al., 2001; Hodges & Patterson, 1995): (1) CF summed across the categories *animals, birds, dogs, household objects, vehicles, and boats*; (2) confrontation naming (N) for subsets of the line drawings of common objects in the Snodgrass and Vanderwort corpus (set size = 48 items in the earlier version of the battery and 64 items in the more recent version, as discussed below); (3) WPM for the same items as those in the naming task, using a within-category 10-alternative forced-choice paradigm; and (4) and (5) the word and picture versions of the Pyramids and Palm Trees test of associative matching (PPTw and PPTp; Howard & Patterson, 1992). Because of changes in the standard testing protocol over the years of data collection, only a subset of the patients routinely performed PPTw (26% of controls, 76% of SD, 69% of PNFA, 58% of fvFTD, 47% of AD, and 50% of PCA cases).

These tasks include two measures of verbal production (CF and N) and three measures of comprehension (WPM, PPTp, PPTw). They also vary in the extent to which each is likely to rely on visual, linguistic, and executive resources in addition to semantic memory. For example, CF does not tax visual processing but requires speech production and executive functions such as working memory, maintenance of task context, and inhibition of prepotent responses. By contrast, WPM and picture naming recruit both visual and linguistic faculties, and the picture version of the PPT draws on visual resources but does not overtly require linguistic processing. It is important to emphasize that, because the five syndromes under investigation are all progressive, performance will be further influenced by stage of disease severity.

In addition to the magnitude of semantic impairment, at least four other factors, by hypothesis, might contribute to impaired performance in each individual semantic task; additional tests were conducted to assess these factors.

Impairment to executive function. The majority of patients were assessed on LF, a task that is particularly sensitive to frontal-lobe damage (e.g., Coslett, Bowers, Verfaellie, & Heilman, 1991). LF imposes executive

Table 1
Details of the Subject Groups and Test Rounds

	Controls	SD	PNFA	fvFTD	AD	PCA
Number/342	106	42	23	40	113	18
Age (years)	70.5 (8.7)	62.6 (6.1)	66.3 (7.4)	58.6 (7.3)	67.3 (8.3)	59.5 (7.1)
Sex (% F)	61	43	39	26	55	40
Education (years)	10.7 (2.1)	10.4 (1.7)	10.7 (1.4)	10.4 (1.3)	11.1 (2.5)	12.1 (2.3)
Examinations/848						
• Total number	106	132	76	60	412	62
• Mean	1	3.1	3.3	1.6	3.6	3.4
• Max/subject	1	16	6	4	13	7

Note. Values in parentheses are standard deviations. AD = Alzheimer's disease; fvFTD = frontal variant frontotemporal dementia; PCA = posterior cortical atrophy; PNFA = progressive nonfluent aphasia; SD = semantic dementia.

demands similar to those of CF but depends less on the integrity of the semantic system (Monsch et al., 1992; Salmon et al., 1999). Participants were assessed on the letters *F*, *A*, and *S*, and data were summed across these.

Visual processing impairments. Four tests from the Visual Object and Space Perception Battery (VOSP; Warrington & James, 1991) were used to evaluate the integrity of visual-processing abilities: (1) the screening test, which is used to determine whether patients have any serious low-level visual deficit; (2) the incomplete letters test, which requires participants to identify capital letters from a degraded, noisy representation; (3) the object decision test, in which participants must determine which in an array of four silhouettes is a real object; and (4) the silhouette-naming task, which requires participants to name drawings of familiar objects from solid silhouette shapes and hence requires semantic, verbal, and visual resources. The VOSP entered our standard test battery partway through the data collection period, and results from it were therefore not available for all participants. The subtests chosen were those most consistently administered across the patient cohorts.

Verbal production deficits. Forward digit span was chosen as a measure of verbal production because it requires fluent verbal production but little if any semantic processing. However, it should be noted that this task also depends on working memory.

Overall severity of cognitive impairment. In multifactor dementias, it is possible that conjoint cognitive deficits may combine to produce a global, nonspecific impairment to all aspects of cognition, resulting, for example, in performance that is generally slow and unreliable relative to healthy controls. Past assessments of progressive syndromes have revealed that a single factor corresponding intuitively to overall disease severity accounts for a large proportion of the variation in performance across a battery of standard cognitive measures (Lambon Ralph, Patterson, Graham, Dawson, & Hodges, 2003). The Mini-Mental State Examination (MMSE; Cockrell & Folstein, 1988) has been shown to correlate highly with this "severity" factor and was used in the current study as a measure of global cognitive impairment.

These specific measures were not used for diagnostic purposes; however, because the patient groups are defined in part on the basis of behavioral criteria, some inevitable overlap occurs between the diagnostic criteria and the neuropsychological measures of interest in our analysis. For instance, SD is defined as a progressive loss of conceptual knowledge in the context of fluent speech and preserved nonverbal episodic memory, a definition that entails anomia and poor verbal comprehension. Although the experimental naming tests and verbal comprehension measures described previously are not the same naming and comprehension tests used for diagnosis, it is not surprising to find that individuals with SD perform poorly on these measures. However, the interest of the following analyses lies in the comparison of performance between SD and the other subject groups on the same semantic measures, taking SD as the canonical example of a semantic syndrome. Because impairment on semantic tasks is not diagnostic for these other patient groups, this comparison allows us to determine (a) whether impairments in other patient groups are of comparable magnitude to those observed in SD, (b) whether impairments arising from semantic or nonsemantic factors can be discriminated in a given task, and (c) whether groups that are matched for severity on a diagnostic semantic task differ in other important respects.

Details relating to the corpus of data. The data that form the basis of our analyses were collected over almost a decade as part of longitudinal studies designed on the case series approach to neuropsychology. As far as possible, all patients were tested with the same materials and were reassessed on a regular basis to enable documentation of progression over time. This methodology has produced a large corpus of data from a variety of patient groups, allowing us to investigate both the central tendencies and the ranges of impairment across different syndromes and tests. Partly as a

consequence of the practical demands and difficulties of testing over long spans of time, the raw data record is not perfectly complete in the following respects.

First, patients who reached floor performance on a given task during one test round were rarely retested on the task in further rounds. Because the syndromes under consideration are all progressive, it was assumed that performance was unlikely to recover significantly from floor performance and therefore that neither the patient nor the research program would benefit from further effortful and time-consuming assessment with such tasks. For the current study, we identified all test rounds in which missing values appeared after floor performance was observed on a given test and replaced these values with floor-level scores (i.e., zero for productive tasks such as naming; chance-level performance for forced-choice tasks such as WPM).

Second, two of the five semantic tasks (WPM, N) were drawn from the Cambridge Semantic Battery first described by Hodges, Salmon, & Butters (1992), subsequently by Hodges et al. (1992), and then updated by Garrard et al. (1998, 2001; see also Bozeat et al., 2000). The older and newer versions of the battery comprise the same sets of tasks constructed around the same broad semantic domains; but the particular items appearing in the two versions are not identical. Specifically, the original battery consisted of 48 items drawn from three categories of living things (land animals, water creatures, and birds) and three of nonliving things (household objects, tools, and musical instruments). The newer version consists of 64 items, drawn from three overlapping but not identical living categories (land animals, birds, and fruits), and three overlapping but not identical nonliving categories (household objects, tools, and vehicles). Approximately two thirds of the measurements in the database were taken with the older version of the tests, whereas the remaining third were taken with the more recent version. Because both versions effectively measure the same underlying cognitive abilities, using the same methods and overlapping items, we combined observations from older and newer versions of each test by converting them to proportions correct and thereafter treating them as the same measure.

Cross-Sectional Versus Longitudinal Data Analysis

Most of the individuals participating in the study were assessed on more than one occasion; multiple observations gathered from the same individual are not statistically independent. Thus in the preliminary analyses, where sufficient cases are found in each group, we report the results of cross-sectional analyses, taking a single data point on each measure from the first testing session for each individual. Where power becomes a concern in the cross-sectional case, we have employed mixed linear models to investigate a larger number of observations from the same group of individuals over time. Mixed linear models permit assignment of within-individual covariation (across successive observations) to the individual, rather than to the patient group that the individual exemplifies. Such models thus allow us to assess differences between patient groups using longitudinal data, after taking into account within-individual covariation.

Normalizing Transformations

An initial consideration of the distributions for the various measures across all participants indicated that several tests deviated strongly from normal. All measures were therefore subjected to appropriate normalizing transformations. Specifically, measures with many ceiling values were converted to proportions and subjected to an arcsin transformation to increase the spread at the top end of the range (MMSE, N, WPM, PPTp, PPTw, VOSP-Scr, VOSP-Inc, VOSP-OD); and measures with substantial degrees of skew were subjected to power transformations (DS, CF, LF, N, WPM, PPTp, PPTw, VOSP-Scr, VOSP-OD). Where arcsin and power transformations were both applied, the measure was first converted to a proportion, then subjected to the arcsin transformation, and finally to the

power transformation. These transformations succeeded in producing Shapiro-Wilks W values above 0.90 for all but one of the transformed measures. The exception was WPM scores, which showed a strong leftward skew with many ceiling values so that the transformed data were difficult to normalize within acceptable limits ($W = 0.81$). All reported significance tests were calculated on transformed data, whereas means and confidence intervals were first calculated for transformed data and were then reverse-transformed to the original range to yield meaningful values in the data presented in the article. It should be noted that raw means (from untransformed data) and confidence intervals for all patient groups on all measures can be viewed in the online supporting materials for this article, together with means and confidence intervals from reverse-transformed data shown in the following figures.

Analysis 1: Comparison of Group Central Tendencies

Demographic Data

Means and standard deviations for age at time of first test, years of education, and sex are shown in Table 1. Univariate analyses of variance (ANOVA) revealed significant differences in age across the groups, $F(5, 334) = 15, p < .001$, with Bonferroni-corrected post hoc contrasts showing that the healthy controls were significantly older than the SD, FTD, and PCA groups. No patient group was older than the healthy controls, so the patterns of impairment apparent in the analyses to come do not likely

reflect a confound with age. No difference was found in the mean years of education across groups ($F(5, 287) = 1.5, p = ns$). The balance of sexes differed across groups, $F(5, 338) = 2.5, p < .05$, but corrected post hoc contrasts revealed that this effect was carried by the two most extreme groups: there were significantly fewer males in the healthy control group compared with the FTD group ($p < .03$); otherwise the groups did not differ reliably.

Nonsemantic Tasks

We begin by examining performance on the seven nonsemantic tests to verify that the patient groups showed the expected patterns of deficits on each. In these analyses we consider cross-sectional group means plus confidence intervals, taking data from the first testing round for each patient; these are shown in Figure 1. Means for different patient groups were compared using univariate ANOVA, and Bonferroni-corrected contrasts were calculated for all group pairings (see the online supporting materials for estimates of effect size in these analyses). The results may be summarized as follows:

MMSE. All groups had significantly lower scores than controls in the MMSE, a measure of global cognitive impairment ($F(5, 278) = 39.8, p < .001$; $p < .001$ for all contrasts to control group). As is apparent in Figure 1, the scores of patients with SD, PNFA, AD, and PCA were comparable on this measure and did not reliably differ from one another, whereas patients with fvFTD scored significantly higher than the other

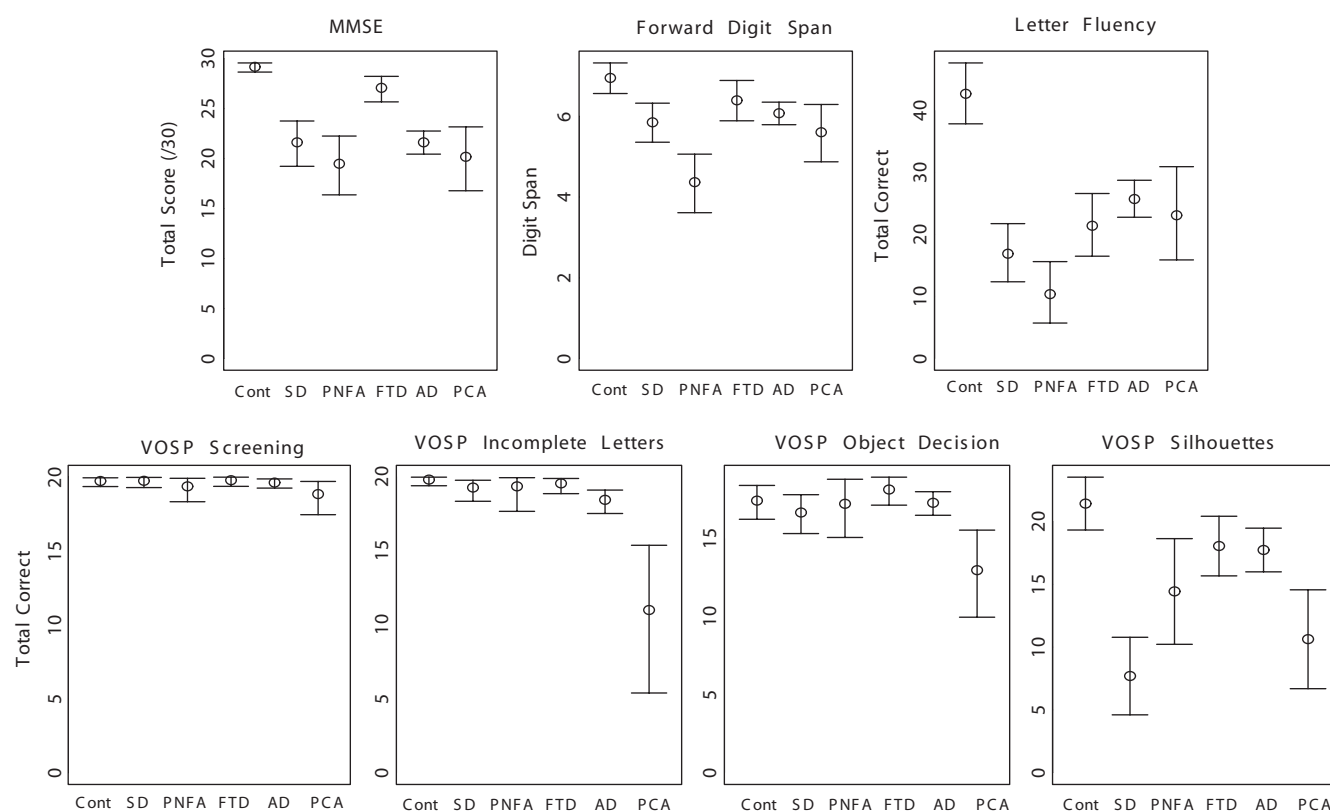


Figure 1. Confidence interval (95%) of the mean for controls and subject groups for each nonsemantic task. The data are cross-sectional: A single observation was taken from each individual at the first available testing round for each measure. AD = Alzheimer's disease; CONT = controls; FTD = frontotemporal lobar degeneration; MMSE = Mini-mental State examination; PCA = posterior cortical atrophy; PNFA = progressive nonfluent aphasia; SD = semantic dementia; VOSP = Visual Object Space Perception Battery.

groups ($p < .001$ for all contrasts). It is worth noting that poor performance on the MMSE, even on the same items, may occur for somewhat different reasons in the various groups. For example, one item in the MMSE requires participants to recall the names of three objects that they heard and repeated a few minutes previously. Individuals with AD may fail this item because of their episodic memory deficits, whereas individuals with PNFA may fail it because of their speech production difficulties. Nevertheless, the MMSE correlates well with overall severity of cognitive impairment regardless of the type of dementia and, from this measure, the SD, PNFA, AD and PCA groups in our sample had similar degrees of global impairment, whereas the patients with fvFTD were more mildly impaired.

Digit span. Significant differences were observed across groups, $F(5, 246) = 9.8, p < .001$. The PNFA group, as expected, was substantially impaired on the forward digit span task and was significantly worse than each of the other patient groups ($p < .001$ for all contrasts), who did not differ significantly from one another. SD, AD, and PCA groups were mildly but significantly impaired on the task relative to controls; the fvFTD group did not differ reliably either from the control group or from the SD, AD, or PCA groups.

LF task. Again, reliable differences were observed across groups, $F(5, 246) = 19.1, p < .001$. All patient groups were characterized by large deficits relative to controls in the total number of correct responses on LF ($p < .001$ for all contrasts). The PNFA group showed the worst performance and the AD group the best, and this specific contrast was the only significant one among patient groups ($AD > PNFA, p < .001$).

VOSP. All groups performed well on the screening component of the VOSP, indicating no serious low-level visual problems ($F(5, 114) = 1.0, p = ns$). On two of the remaining subtests, reliable differences across groups were observed (incomplete letters: $F(5, 142) = 3.4, p < .01$; object decision: $F(5, 140) = 8.3, p < .001$), with post hoc contrasts showing that only the PCA group performed significantly worse than control participants ($p < .001$). The groups also differed on the silhouette naming task, $F(5, 119) = 13.2, p < .001$, with all groups except fvFTD impaired relative to controls ($p < .001$ for all contrasts), reflecting the fact that the test depends on visual, semantic, and verbal production faculties. The PCA and SD groups were reliably more impaired than fvFTD and AD on this task ($p < .001$ for all contrasts). The PNFA patients did not differ reliably from other patient groups.

Summary for nonsemantic tasks. As expected, the digit span task pinpointed a pronounced deficit in PNFA, and the visual subtests of the VOSP likewise marked out the PCA group as impaired relative to both the control subjects and the other patient cohorts. The LF task is clearly sensitive to frontal-executive impairment, because this was the only task (apart from MMSE) on which the fvFTD group differed from controls. LF is also susceptible to other forms of cognitive impairment: all groups, especially PNFA, revealed serious deficits in the task. Notably, the performance of the SD group was not significantly worse than that of the other patient groups on any of these measures, the sole exception being silhouette naming in the VOSP, a task with a clear semantic component.

Overview of Intergroup Comparisons for the Semantic Tasks

Figure 2 shows the confidence intervals of the cross-sectional means in each group for the five semantic tests. As before, group means were compared using univariate ANOVA, and all pairwise contrasts were calculated. Effect sizes appear with the online supporting materials for this article. The results may be summarized as follows:

Naming. Significant differences across groups were observed, $F(5, 325) = 49.5, p < .001$, with all subject groups impaired relative to controls ($p < .001$ for all contrasts). AD, PCA, and fvFTD were only mildly impaired and did not differ significantly from one another. PNFA patients showed somewhat larger deficits (significantly worse than fvFTD and AD, $p < .001$ for both contrasts); and SD patients were most impaired (significantly worse than all groups except PNFA, $p < .001$ for all contrasts).

CF. Significant differences were observed across groups, $F(5, 325) = 60.3, p < .001$, and all groups were impaired at this task relative to controls ($p < .001$ for all contrasts). The SD and PNFA groups had comparably poor levels of performance, with fvFTD, AD, and PCA groups performing somewhat better ($p < .001$ for all contrasts) and not significantly different from one another.

WPM. Reliable differences were observed across groups, $F(5, 322) = 33.1, p < .001$. Owing to the ceiling effect in this task for controls, all patient groups revealed a significant impairment relative to controls ($p < .004$ for all contrasts); but deficits were mild for all patient groups except SD and did not differ significantly from one another. The SD group was significantly more impaired than all other groups ($p < .001$ for all contrasts).

PPT-pictures. Reliable differences were observed across groups, $F(5, 236) = 30.2, p < .001$. Again, normal performance is at ceiling on this task, and hence all patient groups were impaired relative to controls ($p < .001$ for all contrasts). PNFA, fvFTD, and AD groups showed equivalent and mild degrees of impairment, whereas SD performance was significantly worse than all subject groups except PCA ($p < .01$ for all contrasts), which did not differ significantly from any other patient group.

PPT-words. Reliable differences were observed across groups ($F(5, 157) = 18.0, p < .001$). All groups were impaired relative to controls ($p < .05$ for all contrasts) except PCA. Four of the five subject groups (PNFA, fvFTD, AD, PCA) did not differ from one another; the remaining group (SD) performed significantly worse than all other cohorts except PNFA ($p < .001$ for all contrasts).

Summary of semantic tasks. Two general observations are of interest. First, all subject groups except PCA revealed a statistically reliable degree of impairment on all of the semantics tasks; and even the PCA group was significantly impaired on four of the five tasks. The data thus suggest either that all subject groups have at least mild semantic impairments or, more likely, that the semantic tasks themselves are sensitive to other kinds of cognitive impairment. Second, the SD cohort was characterized by a greater degree of impairment on all semantic tasks than the other groups. On the three tests of comprehension, the contrasting subject groups did not differ significantly from one another, and all performed significantly better than the SD group, with two exceptions: the PCA group, which did not differ significantly from the SD group on the *visual* comprehension task (PPT-pictures), and the PNFA group, which did not differ significantly from the SD group in the *verbal* comprehension task (PPT-words). On the two production tasks, the SD and PNFA groups did not differ; otherwise the SD group showed substantially greater deficits than the other groups.

Summary of Analysis 1

In summary, all patient groups were significantly impaired relative to controls on virtually all semantic tasks (the single exception being the PCA group on PPTw); but the magnitude of impairment was significantly greater for the SD group compared with all other patient groups. In contrast, the SD group was not more impaired than other groups on the nonsemantic tasks. In general, the pattern of performance was consistent with the syndromes in question: the PNFA group was impaired relative to other groups on all tasks requiring verbal production; the fvFTD group performed generally well on all tasks except the two verbal fluency tasks, which are known to tax executive function; and the PCA group performed well relative to other groups on all tasks save those that tax visual perception. The AD group showed a mixed pattern of impairment but, surprisingly, showed no tendency toward worse scores on semantic tasks than other groups not typically attributed semantic deficits. The test often called on to reveal semantic impairment in AD (CF) posed great difficulty for all patient groups, reflecting its multicomponent nature. Therefore this initial comparison provides no real basis to conclude that subnormal performance on CF or other semantic tasks in AD primarily reflects an underlying semantic deficit. Analysis 1 also makes it clear that each semantic task may be sensitive to different kinds of cognitive impairment,

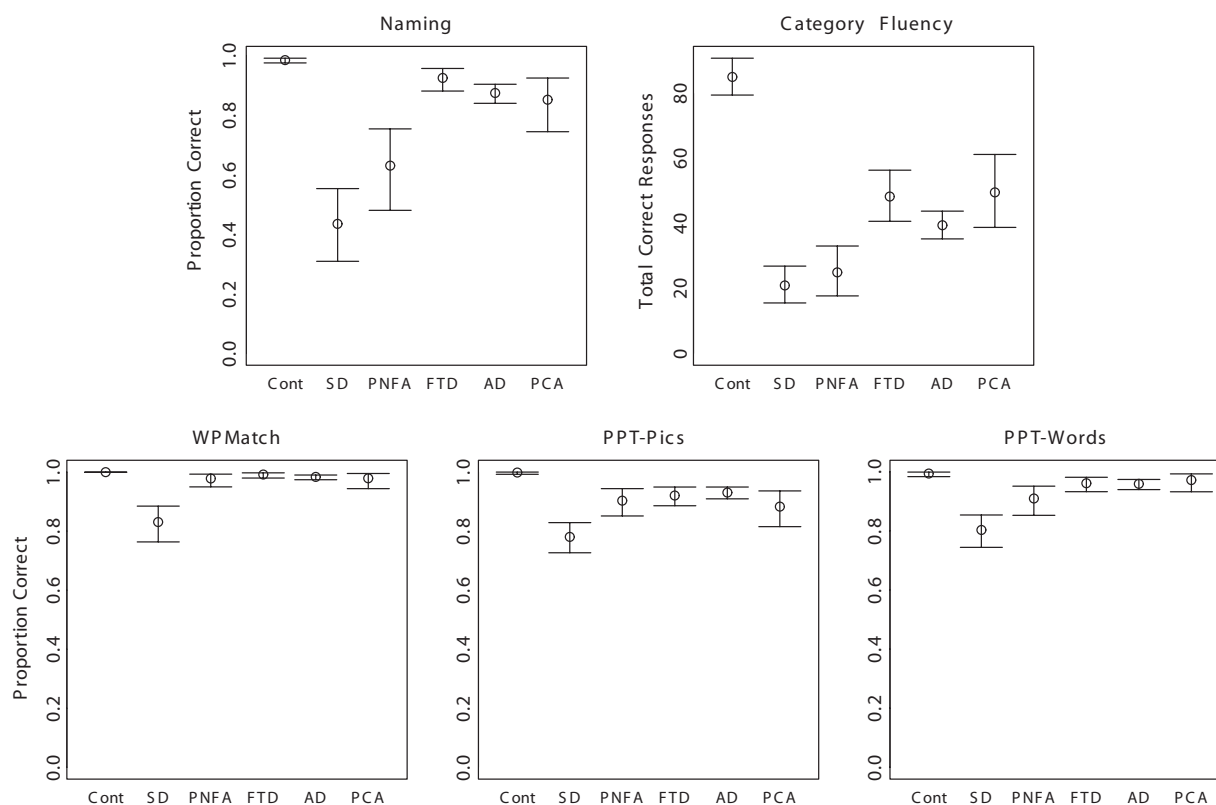


Figure 2. Confidence interval (95%) of the mean for controls and subject groups for each semantic task. The data are cross-sectional: A single observation was taken from each individual at the first available testing round for each measure. AD = Alzheimer's disease; CONT = controls; FTD = frontotemporal lobar degeneration; PCA = posterior cortical atrophy; PNFA = progressive nonfluent aphasia; PPT-Pics = Pyramids and Palm Trees Test-Picture version; PPT-Words = Pyramids and Palm Trees Test-Word version; SD = semantic dementia; WP Match = word-picture matching test.

including executive dysfunction, visual impairment, and verbal production deficits. To determine whether impairments to these faculties coexist with a semantic deficit in the various syndromes, it is necessary to tease these factors apart in the different tests.

Analysis 2: Contrasting Pairs of Semantic Tasks

Executive and Semantic Contributions to Verbal Fluency

In Analysis 1, all patient groups showed serious impairment to the two verbal fluency measures. Although both LF and CF likely depend to some extent on semantic processing, CF is typically considered to tax the semantic system more heavily (Monsch et al., 1992; Salmon et al., 1999). To determine whether performance deficits primarily reflect semantic or executive factors, we therefore compared relative performance on the two different tasks.

We began by contrasting all patient groups regardless of severity, then considered subsets of groups classified by the magnitude of their comprehension impairment (as assessed by WPM) into Mild, Moderate, and Severe. The latter contrasts allow us to determine, for groups of patients with equivalently poor performance on a common test of semantic impairment, how performance compares on LF and CF. If group differences remain robust in these contrasts, then they do not likely arise from the overall magnitude of semantic impairment but rather from qualitative

differences in the syndromes. For the initial contrasts, cross-sectional data were employed. For the comparison by severity, because of the comparatively large number of cells in the analysis (5 patient groups \times 3 severity groups), longitudinal data analyzed with mixed linear models were employed. As in the previous analysis, effect size estimates and model parameter estimates may be viewed with the online supporting materials for this article.

Methods for analysis independent of severity. In this contrast of LF and CF, we consider the total number of correct responses provided by each participant, summed across the different fluency categories in each condition (i.e., the different semantic categories in CF; the different letters in LF). For the particular categories used in our tests, normal control subjects typically generate a greater number of responses for CF than for LF. To contrast the magnitude of patient deficits, the patient data from the two tasks were therefore normalized to the same scale by dividing each individual's score by the control mean for the corresponding task. Thus each individual data point reflects the number of correct responses given by a single patient, as a proportion of the average number of correct responses typically provided by roughly age-matched controls. These measures were transformed to normal by taking the square root of the natural log ($W > 0.92$). As before, significance tests were calculated for transformed data; means and confidence intervals were calculated in the transformed space and reverse-transformed to generate the figures in this section.

Performance on CF and LF was contrasted using repeated measures ANOVA, with task type (LF or CF) as the within-subjects measure and patient group as a between-subjects factor. Where interactions between group and task type were observed, we investigated these by calculating the difference between CF and LF for each individual and analyzing the resulting data with univariate ANOVA.

Results independent of severity. The means for the different patient groups are shown in the left panel of Figure 3. The repeated measures ANOVA revealed a main effect of patient group, $F(1, 191) = 5.56, p < .001$, because some patient groups (most notably, fvFTD, AD, and PCA) produced substantially more responses than others (SD and PNFA). The main effect of task type was not significant, but there was a strong interaction between task type and patient group, $F(4, 180) = 7.28, p < .001$. The source of the interaction is apparent from the plot of means: three of the patient groups (PNFA, fvFTD, PCA) showed little difference in performance between LF and CF tasks, whereas the SD and AD groups produced considerably fewer correct responses (as a proportion of control means) in CF than LF. To identify the patient groups for which the difference between CF and LF was statistically reliable, we calculated the 95% confidence interval of the difference between CF and LF for each group. The confidence intervals included zero for PNFA, fvFTD, and PCA groups, indicating no reliable difference between task types. For SD and AD, the mean difference was reliably negative (SD = -0.20 to -0.04 ; AD = -0.17 to -0.08), indicating LF > CF.

Methods for patients grouped by severity of semantic impairment. To determine whether this pattern is observed across the spectrum of disease severity, longitudinal patient data were analyzed with a mixed linear model. Specifically, the difference between CF and LF was calculated for all testing sessions with each individual patient. For the PCA and fvFTD groups, only a small number of patients had more than four observations; thus we included only the first four observations per individual for all patient groups in the analysis. These data were then entered into a mixed linear model in which the difference between fluency measures (CF – LF) was the dependent measure; patient group was treated as a fixed, between-subjects factor; and testing round was treated as an ordinal index of the repeated effect within individual. Thus in this analysis, we were able to

incorporate observations from patients in earlier and later phases of their disease progression, rather than focusing solely on their performance at the time of first testing.

Mixed models permit the theorist to specify an expected covariance matrix for the repeated effect, if the within-individual covariance structure is known, or to estimate the within-individual covariance from the data. The former approach is preferable because the fitting procedure requires fewer degrees of freedom when the covariance structure is specified. We therefore adopted the following strategy in this and subsequent analyses: We first fitted the model with an unstructured covariance matrix, requiring the within-individual covariance to be estimated from the data. We then fitted the model with a compound-symmetry matrix specified. The compound-symmetry matrix assumes constant variance for each testing session and equal covariance across all pairs of testing sessions. To determine whether the compound-symmetry matrix reflects accurate assumptions about the structure of our data, we compared the fit of the unstructured model to that of the compound-symmetry model. In all cases, parsimony-adjusted measures of model-fit (specifically, the Akaike Information Criterion [AIC]; Bozdogan's (1987) corrected form of the AIC; and Schwartz's (1978) Bayesian Information Criterion) were comparably good or better for the compound-symmetry model—hence we report results from the analyses with a compound-symmetry covariance structure specified.

To determine whether the difference between CF and LF varies as a function of severity of semantic impairment, all patients were classified according to the presence/magnitude of their comprehension deficits (as assessed by WPM) as Mild (WPM scores from 61 to 64), Moderate (56 to 60), and Severe (55 or less). This factor was then added to the mixed linear model and treated as a fixed, between-case factor. It should be noted that, in Analysis 1, some patient groups performed near ceiling on this measure, with little variation. This analysis only incorporated observations from the first round of testing for each patient, however. In all patient groups, a greater spread of performance is observed when measures from subsequent testing sessions are included in the data. Thus in all analyses that group patients by the severity of their comprehension impairment, all patient groups are well represented at each level of severity.

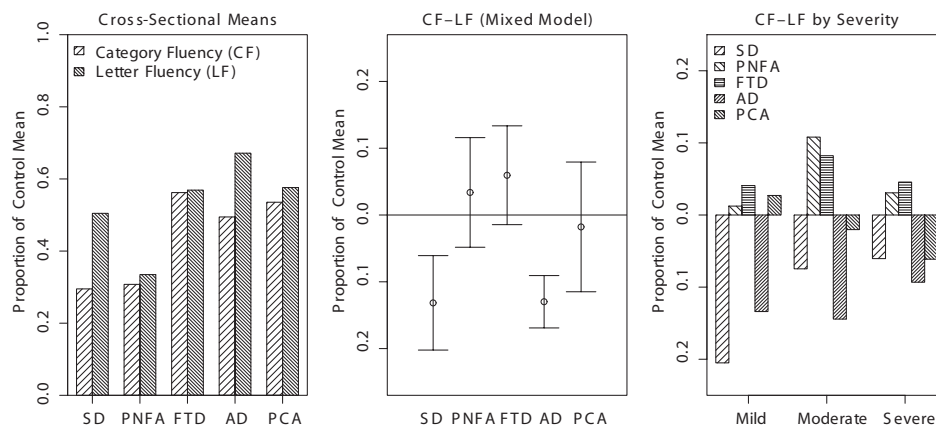


Figure 3. Contrast of category fluency (CF) and letter fluency (LF) for each subject group. *Left:* Number of correct responses produced in both tasks, as a proportion of the control mean for the same task. Data are cross-sectional, taking the first available observation from each subject. *Middle:* 95% confidence interval for the expected value of the difference CF – LF for the five subject groups from the corresponding mixed linear model in Analysis 2. The data incorporate the first four observations available from each subject. *Right:* Expected values for the difference CF – LF in subject groups with mild, moderate, or severe comprehension impairments, when severity is added to the model. Data are longitudinal, taking the first four available data points from each subject. AD = Alzheimer's disease; CONT = controls; FTD = frontotemporal lobar degeneration; PCA = posterior cortical atrophy; PNFA = progressive nonfluent aphasia; SD = semantic dementia.

Results

Complete observations were available for 188 patients in total, including 29 SD, 29 fvFTD, 22 PNFA, 95 AD, and 13 PCA. An average of 2.3 observations from each patient was included in the analysis. The mixed model predicting the discrepancy between CF and LF from patient group alone revealed reliable differences across patient groups, $F(5, 199) = 11.9, p < .001$. Confidence intervals for the expected values within each group are shown in the middle panel of Figure 3. Just as in the cross-sectional analysis, patients in the SD and AD group performed comparatively worse on the CF than the LF task, so that the expected values for the difference CF – LF is reliably negative in these groups. For all other groups, the confidence intervals for the expected difference overlap with zero, indicating no reliable difference between tasks for these groups.

When severity of the comprehension impairment was added to the model, the main effect of patient group remained robust, $F(4, 203) = 7.1, p < .001$. No effect of severity was observed ($F(2, 385) = 0.3, p = ns$), nor was the interaction of severity with patient group reliable ($F(8, 370) = 0.7, p = ns$). Thus the propensity for patients in the SD and AD groups to perform worse on CF than LF appears to be consistently observed across the range of severity; and patients in other groups appear to perform equivalently in the two tasks, regardless of their performance in the comprehension task. The right panel of Figure 3 shows the expected values of the difference CF – LF for each patient group, plotted separately for those with Mild, Moderate, and Severe comprehension impairments. In all cases, the difference for the SD and AD groups is negative, whereas the difference for the FTD and PNFA groups is positive. Although patients in the Severe PCA group appear to be doing somewhat worse in the CF relative to the LF task, this discrepancy is not statistically reliable.

Summary of results for the contrast of CF and LF. The comparison of LF and CF revealed two distinct patterns in impaired fluency across patient groups. Three groups (PNFA, fvFTD, and PCA) showed equivalent degrees of impairment in the two tasks, regardless of the magnitude of comprehension deficits. The data suggest that the aspects of cognitive dysfunction that characterize these disorders, though they may greatly impair verbal fluency generally, will influence CF and LF about equally. By contrast, semantic impairment clearly impacts CF more than LF: The SD group, known a priori to have semantic deficits, generated fewer correct responses on the CF than the LF task, regardless of the severity of their comprehension deficits; the AD group showed a remarkably similar pattern, suggesting that verbal fluency impairments in AD are at least partly, and perhaps predominantly, due to an underlying semantic impairment, rather than to executive or verbal production deficits.

Contrasting Comprehension of Pictures Versus Words

As stated previously, Analysis 1 revealed that the PCA group performed as poorly as patients with SD on a purely visual test of comprehension (PPTp), better than patients with SD but worse than controls on WPM, and with no reliable deficit relative to controls on a purely verbal test of comprehension (PPTw). The data suggest that tests such as WPM and PPTp are susceptible to visual impairment when comprehension is relatively spared. Might poor performance on WPM or PPTp also be attributable, at least in

part, to underlying visual impairment in other patient groups? To investigate this issue, we contrasted performance on the word and picture versions of the PPT for all five patient groups, using methods exactly analogous to the contrast of CF and LF in the previous section.

Results independent of severity. The repeated-measures ANOVA revealed a main effect of patient group, $F(4, 125) = 11.23, p < .001$, indicating as shown previously that the SD group is reliably worse than other patient groups across both tasks ($p < .001$ for all Bonferroni-corrected contrasts except to PNFA; $p < .05$ for PNFA). Although the left-hand panel of Figure 4 shows very little difference in performance across the two tasks for most groups, the main effect of task type was also statistically reliable, $F(1, 125) = 5.424, p < .03$, with patients' scores slightly better on the word than the picture version of the task overall. The interaction between patient group and task type was not statistically reliable ($F(4, 125) = 1.88, p = ns$).

Part of the reason for this null result may be the comparatively small number of participants in some of the patient groups in this analysis. As noted previously, PPTw was only administered routinely to a subset of the complete cohort; hence limited data were produced for this contrast, especially for the group of particular interest (31 SD, 16 PNFA, 23 fvFTD, 51 AD, but only 9 PCA). To determine whether the null result arose from a lack of statistical power, we again analyzed longitudinal data in a mixed linear model predicting the difference between PPTw and PPTp from patient group, and using testing round as the index of the repeated effect. As in the previous analysis of verbal fluency, only the first four observations from each participant were used in the analysis. Valid measures were available for 127 patients total, including 31 SD, 21 FTD, 16 PNFA, 52 AD, and 7 PCA. An average of two observations per patient were included in the model; and of particular interest, 20 observations total were available from the PCA group. The model revealed a reliable main effect of patient group, $F(5, 136) = 3.7, p < .005$. The middle panel of Figure 4 shows the confidence intervals of the expected value in each patient group for the difference PPTw – PPTp. The expected value for the PCA group is reliably positive, indicating overall better performance in this group for words compared with pictures. All other confidence intervals overlap with zero.

Results divided by severity of semantic impairment. To determine how the contrast of PPTw and PPTp varies with the severity of comprehension impairment, patients were again classified as Mild, Moderate, or Severe on the basis of their WPM scores, and this factor was added to the mixed linear model. In this model, patient group remained a significant predictor of the difference PPTw – PPTp, $F(4, 142) = 5.7, p < .001$. A reliable main effect of the severity of comprehension impairment was also found, $F(2, 224) = 6.8, p < .001$, as well as a reliable interaction between patient group and severity, $F(8, 220) = 4.0, p < .001$. Expected values for each patient group, split by severity, are shown in the right panel of Figure 4. For the PCA group, the discrepancy between word and picture comprehension grows as performance on the comprehension measure declines. Two of the patient groups (SD and PNFA) show no difference between PPTw and PPTp, regardless of the overall severity of their comprehension impairments. The remaining two groups (AD and fvFTD), somewhat surprisingly, show no discrepancy between the two tasks when comprehension impairments are mild or moderate but, like the

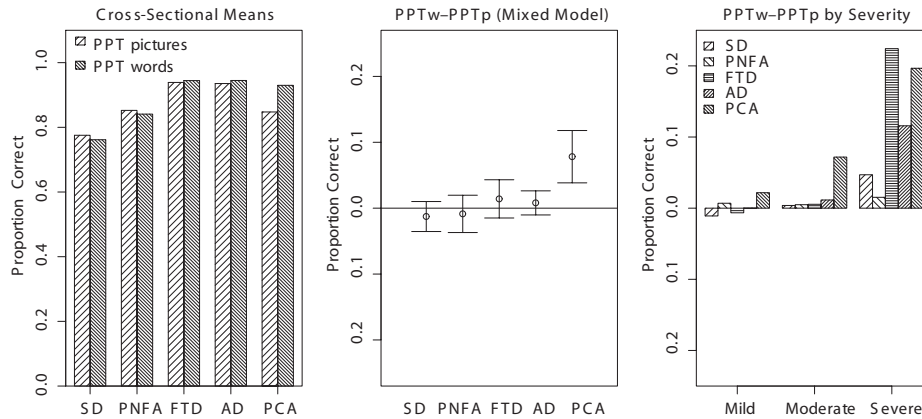


Figure 4. Contrast of verbal and visual comprehension (purely verbal test of comprehension [PPTw] and purely visual test of comprehension [PPTp]) for each subject group. *Left:* Proportion correct for both tasks, from cross-sectional data. *Middle:* 95% confidence intervals for the expected value of the difference PPTw – PPTp for each subject group from the corresponding mixed linear model in Analysis 2. The data incorporate the first four observations available for each subject. *Right:* Expected values for the difference PPTw – PPTp, for subjects with mild, moderate, and severe comprehension impairments, when severity is added to the model. AD = Alzheimer's disease; CONT = controls; FTD = frontotemporal lobar degeneration; PCA = posterior cortical atrophy; PNFA = progressive nonfluent aphasia; PPTw–PPTp = Pyramids and Palm Trees Test words—Pyramids and Palm Trees Test picture; SD = semantic dementia.

PCA group, perform considerably better on the PPTw task than the PPTp task when comprehension impairments are severe.

Summary of contrast for picture versus word comprehension. In summary, only patients with PCA showed a discrepancy between word and picture comprehension that increased with the magnitude of the comprehension impairment measured by WPM. The SD group showed equivalent performance for both tasks across all levels of severity, suggesting (a) that their comprehension problems do not stem from visual-processing impairments as in PCA, and (b) that the semantic impairment in SD compromises visual and verbal comprehension equally. Neither the PNFA group nor the Mild and Moderate cohorts of the AD and fvFTD groups showed any discrepancy between visual and verbal comprehension—again suggesting that visual-processing impairments do not contribute significantly to comprehension deficits in these groups.

The Severe cohorts of the AD and fvFTD groups did show a marked PPTw > PPTp advantage. In the case of fvFTD, this discrepancy is unlikely to have the same visual-processing source as the similar effect in PCA, given that the two groups differ in almost every other respect. This outcome suggests that the PPTp may be vulnerable to executive dysfunction, perhaps because such patients have difficulty inhibiting responses to visually complex or interesting stimuli. As a consequence, it is difficult to interpret the analogous advantage for words in AD, which may reflect visual-processing deficits (as in PCA) or the influence of executive dysfunction (as in fvFTD).

Contrasting Production and Comprehension

Analysis 1 revealed that three of the patient groups had only mild naming deficits, concurrent with mild comprehension deficits and more severe impairment of category and LF. In contrast, one group (PNFA) showed severe deficits in both naming and verbal fluency measures, coupled with mild deficits in comprehension;

and one group (SD) showed not only comparably poor production in naming and fluency but also very poor comprehension. Production deficits were even more profound in the SD group than in the PNFA group, leading to the question of whether SD patients have a purely semantic impairment or an additional postsemantic deficit in verbal production. To investigate this question, we contrasted naming and comprehension directly, using the methods by now familiar from the previous two analyses.

Results for contrast independent of severity. Mean proportions correct for all patient groups in the two tasks are shown in the left panel of Figure 5. As expected given differences in task demands, all patient groups were more successful in the comprehension than the naming task, but the difference was especially pronounced for the SD and PNFA groups. The ANOVA revealed a reliable main effect of both patient group, $F(4, 213) = 18.82, p < .001$, and task type, $F(1, 213) = 154.41, p < .001$, as well as a significant interaction between these, $F(4, 213) = 4.6, p < .001$. Bonferroni-corrected contrasts revealed that, across tasks, scores for the SD group were reliably worse than for all other groups ($p < .001$ for all contrasts except to PNFA; $p < .02$ for contrast to PNFA), and that scores for the PNFA group were reliably worse than those of the fvFTD group ($p < .02$). No other group contrasts across tasks were significant.

To compare the average magnitude of the discrepancy between comprehension and production across the different groups, we calculated the difference between WPM and N for each individual and conducted a univariate ANOVA on these data (subjected to a normalizing transformation) using patient group as the fixed independent factor. The results showed a reliable effect of patient group, $F(5, 314) = 33.98, p < .001$, with corrected post hoc contrasts revealing no reliable difference among fvFTD, AD, and PCA groups; a significantly higher WPM > N difference for SD

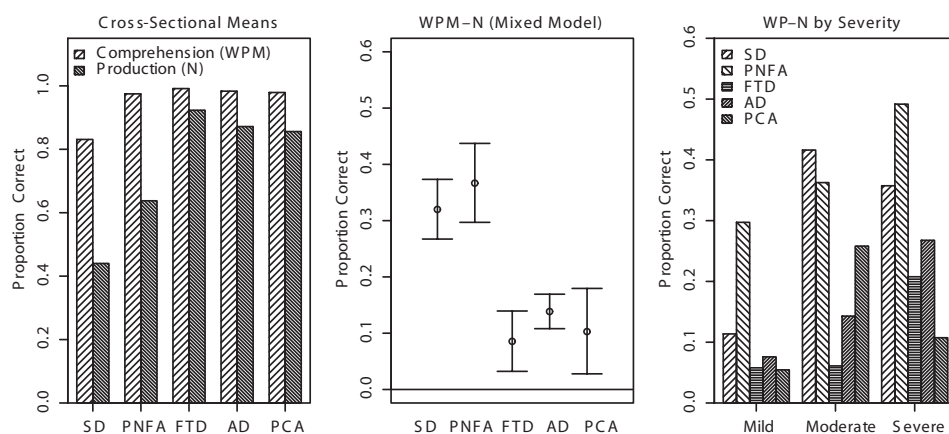


Figure 5. Contrast of production (N) and comprehension (word-picture matching [WPM]) for each patient group. *Left*: Proportion correct for both tasks, from cross-sectional data taking the first available observation for each patient. *Middle*: 95% confidence interval for the expected value of the difference WPM – N from the corresponding mixed linear model in Analysis 2. The data incorporate the first four available observations from each patient. *Right*: Expected values for the difference WPM – N for patients with mild, moderate, and severe comprehension impairments, when severity is added to the model. AD = Alzheimer's disease; CONT = controls; FTD = frontotemporal lobar degeneration; PCA = posterior cortical atrophy; PNFA = progressive nonfluent aphasia; SD = semantic dementia; WP–N = word-picture–matching score (proportion correct) minus naming score (proportion correct).

and PNFA relative to these groups ($p < .001$ for all contrasts); and no difference between SD and PNFA.

Results by severity of comprehension impairment. To contrast comprehension and production, the difference between WPM and N was calculated for the first four observations in which both measures were available in each patient, and these differences were analyzed in a mixed linear model as in the previous analyses. Valid observations were available for 216 patients, including 39 SD, 36 fvFTD, 21 PNFA, 106 AD, and 14 PCA. On average, 2.3 observations were included for each patient, yielding 505 total observations. Patient group was a reliable predictor of the difference WPM – N, $F(5, 223) = 19.8, p < .001$. Confidence intervals of the expected values for each patient group are shown in the middle panel of Figure 5. Although all groups performed reliably better for comprehension compared with naming, this discrepancy was much larger (and equivalently large) for the SD and PNFA groups, confirming that the findings from the cross-sectional data apply to the longitudinal data as well.

When severity of the comprehension impairment is added to the model, patient group remains a reliable predictor of the difference WPM – N, $F(4, 209) = 15.4, p < .001$. The main effect of severity was not a significant contributor to the model ($F(2, 479) = 1.8, p = ns$), but interacted significantly with patient group, $F(8, 428) = 3.8, p < .001$. The right panel of Figure 5 shows expected values for the different patient groups split by severity. Among patients with only mild comprehension impairments, the PNFA group showed a greater discrepancy between naming and comprehension than did the other groups. However, in the moderate and severe cohorts, SD and PNFA groups both had considerably worse performance on naming than on comprehension, relative to the other groups.

Summary of contrasts for production versus comprehension. It is not surprising that all patient groups achieved higher scores on the

comprehension than the production task. More interesting is the contrast in the magnitude of this difference across groups. As expected, PNFA patients were characterized by a larger discrepancy than the other *nonsemantic* groups, regardless of severity, reflecting verbal output deficits that dramatically outstrip their semantic impairments (as found in Analysis 1). The discrepancy between production and comprehension, however, both in the overall patient data and for Moderate and Severe cohorts, was of similar magnitude in SD and PNFA. The data thus suggest that, when comprehension impairments are moderate to severe, patients with SD have greater difficulty producing spoken names than do other patient groups with comparably poor comprehension, excepting only the group known to have a specific deficit of verbal production.

Discussion of Results From Analysis 2

Analysis 2 allows researchers to begin to tease apart the strands contributing to impairment on the five different semantic tasks. From Analysis 1 it was apparent that poor verbal fluency can arise as a consequence of multiple factors, including executive, verbal production, and semantic deficits, making it difficult, for example, to attribute poor performance on CF in AD exclusively to semantic impairment. The contrast of LF and CF in the current analysis, however, revealed two distinct profiles across patient groups. Patients with suspected executive (FTD), verbal production (PNFA), or visual/attentional impairments (PCA) performed equally poorly on both fluency tasks, whereas patients with semantic impairments (SD) were considerably worse on CF than LF. The AD group, although generally somewhat more successful than the patients with SD, showed the same $CF < LF$ pattern—suggesting that their poor CF is mainly attributable to a similar

underlying semantic impairment and not to executive or verbal production disorders. These patterns were apparent across the spectrum of severity.¹

Analysis 1 also indicated that success on tests of comprehension involving arrays of pictures could be compromised by visual-processing deficits, with relative sparing of verbal comprehension. Analysis 2 demonstrated that, although reliable differences between picture and word comprehension may be observed in patients with known visual-processing dysfunction, this factor does not appear to contribute significantly to impaired comprehension in the other patient groups. With the exception of PCA, no patient group showed reliable differences overall between visual and verbal versions of the PPT task. When severity was taken into account, the severe fvFTD subgroup was found to have as large a PPTw > PPTp discrepancy as the severe PCA subgroup, although almost certainly for a different reason given the differences in their cognitive profiles and regions of atrophy. It is therefore difficult to interpret the more modest advantage for PPTw over PPTp in the severe AD group, which may arise from visual, executive, or other factors. Neither the SD nor the PNFA groups showed a discrepancy between words and pictures in the task, however, suggesting that neither semantic nor verbal production deficits will differentially impair comprehension of visual versus verbal materials.

Finally, Analysis 2 established that, although forced-choice comprehension (WPM) is (not surprisingly) somewhat better than verbal production (naming) for the same items in all patient groups, the discrepancy is substantially and equivalently larger in the SD and PNFA groups than in the other three groups for patients with Moderate and Severe comprehension deficits. This outcome is expected for patients with PNFA given their labored and inaccurate speech output, but what explains it in SD? One possibility is that a deficit of semantic processing disproportionately impairs verbal production capabilities. The fluency data suggest, however, that the patients with AD had degraded semantic processing, but they did not have a large discrepancy between production and comprehension. A second possibility is that SD entails both semantic and verbal production deficits, but little evidence exists of such production deficits in nonsemantic tasks.

What accounts for the large discrepancy between production and comprehension observed in SD and PNFA but not in AD? One explanation may be that the typical level of semantic deficit (as measured by comprehension) is considerably more profound in SD than in AD. If verbal production declines nonlinearly with comprehension, then it is possible that patients (of any disease group) with more severe semantic impairment will have a larger discrepancy than milder patients. The contrast of production and comprehension across subgroups divided by degree of comprehension deficit provides partial support for this hypothesis: Among patients with mild comprehension deficits, SD and AD groups did not differ in the discrepancy between WPM and N. Another likely explanation, in our view, derives from the hypothesis and analysis published by Lambon Ralph et al. (2001). As previously mentioned, the temporal-lobe atrophy in SD is usually strikingly asymmetrical, and for the majority of cases in our cohort, the burden of atrophy was greater in the left than the right hemisphere. Lambon Ralph et al. (2001) argued that semantic representations are bilaterally distributed, but phonological representations for speech production are left lateralized. On the general principle that close connections in the brain are more influential than distant ones

(Plaut, 2002), this hypothesis predicts that—even if subgroups of patients with SD with greater left- versus right-sided abnormality have comparable comprehension deficits—the subgroup with predominantly left-temporal atrophy will have a disproportionate deficit in naming and in expressive (i.e., propositional) speech more generally, even if postsemantic mechanisms of verbal production are comparatively spared.

It is worth inquiring whether the previous results depend on our use of the WPM task as a general index of semantic impairment. WPM is a commonly used tool for assessing the degree of semantic impairment (e.g., Caplan, 1987) and thus seems appropriate given our aim of comparing performance across other semantic tasks for groups that are matched for their level of comprehension impairment. Our use of the task does raise two questions, however. First, performance is near ceiling on average for several patient groups (see Analysis 1), so the measure may not provide a very fine instrument for discriminating patients with mild and moderate semantic impairment. Second, in focusing specifically on the severity of comprehension impairments, we may have neglected other important aspects of disease severity. In response to these issues we wish to note that each of the analyses described previously was also conducted using MMSE as the basis for assigning patients to Mild, Moderate, or Severe groups. MMSE is a multi-factor measure sensitive to many different aspects of cognitive dysfunction; and as indicated in Analysis 1, all patient groups were somewhat further from ceiling on this measure. Nevertheless, the results of the analyses discussed previously were qualitatively identical regardless of whether WPM or MMSE was used to scale severity.

Analysis 3: Factors Contributing to Semantic Task Performance

The previous analysis permitted us to deconstruct some of the factors contributing to impoverished semantic task performance in the different syndromes under consideration. In the final analysis we undertook a more formal determination of the factors underlying patient behavior across all the measures from Analysis 1, using exploratory factor analysis techniques. Specifically, we inquired whether the pattern of responses across all tasks and patient groups may be understood as arising from disruption to the three cognitive subsystems assessed in Analysis 2—semantics, visual perception, and verbal production—or by other cognitive factors not yet considered (e.g., executive dysfunction).

Factor Analysis

Methods for factor extraction. Data from 12 measures were subjected to an exploratory factor analysis. These included the four VOSP subtests (screening, incomplete letters, object decision, and silhouette naming), forward digit span, MMSE, LF and CF, nam-

¹ It should be noted that, despite somewhat similar profiles in the current analysis, the PNFA and fvFTD impairments to verbal fluency almost certainly do not arise from identical underlying factors. Although individuals with PNFA have great difficulty producing sequenced speech in any context (see, for example, their forward digit span scores in Analysis 1 and Graham, Patterson, & Hodges, 2004), the same is not true of those with fvFTD.

ing, and word- and picture-versions of the PPT. All available observations for all patients, as described in Analysis 1, were employed. When calculating the correlation matrix, missing values were excluded on a pairwise basis, so that an individual contributed to the correlation of any pair of variables for which that individual had complete data, even if he or she was missing scores on other variables. The distributions of scores on all measures except the VOSP silhouette-naming subtest deviated substantially from normal, and were subjected to the same normalizing transformations described in Analysis 1. The analysis was conducted using the SPSS factor analysis routines. Factors were extracted using principal components analysis, and the results were subjected to a Varimax rotation. In accordance with our hypothesis from the previous analyses that the patient data reflect at least three underlying factors, the algorithm was constrained to yield a minimum of three principal components, and was free to add further components with eigenvalues exceeding 1.

Results for factor extraction. The principal components analysis yielded three factors, with initial (unrotated) eigenvalues of 7.14, 1.26, and 0.89, respectively. The remaining principal components had initial eigenvalues less than 0.73 and hence were not included in the solution. The rotated solution accounted for 77% of the variance in the data set, with factors 1 to 3 contributing 41%, 19%, and 17%, respectively.

Factor loadings for the rotated solution are shown in Table 2. Loadings smaller than 0.30 have been suppressed for clarity. Each factor has a readily interpretable pattern of loadings. Factor 1 loads heavily on all tasks that have a semantic component, including all semantic task measures (WPM, PPTw, PPTp, N, CF), as well as some of the *nonsemantic* tasks that clearly draw in part on semantic resources (VOSP silhouette naming, VOSP object decision, and LF). Factor 2 loads most highly on tasks that require verbal production: DSF, LF, CF, N, and also on the VOSP letter identification subtest. Factor 3 receives loadings from predominantly

visual tasks, including all of the VOSP subtasks, as well as the PPTp. MMSE, known a priori as a general measure sensitive to many different aspects of disrupted cognition, loads highly on the first two factors but not the third. From these patterns it seems clear that the first factor indexes semantic impairment, the second indexes verbal production impairment, and the third indexes disruption to visual processing.

Weightings on the three factors were calculated for every individual with complete data on all measures in the factor analysis. Because there were a large number of cases with at least one missing value, the resulting numbers of observations in each patient group were not large (SD = 10, PNFA = 8, fvFTD = 11, AD = 35, PCA = 6). Nevertheless, a contrast of the patient group weightings using separate univariate ANOVAs for the three factors yielded interesting and statistically reliable differences.

Figure 6 shows the confidence intervals of the mean weightings for the different patient groups on the three factors. Each factor discriminates one group from the overall patient mean: SD cases score much lower on Factor 1, PNFA on Factor 2, and PCA on Factor 3. The PCA group also has somewhat lower weightings on Factor 2. A univariate ANOVA confirms that differences across patient groups are reliable for all three factors ($F(4, 65)$ ranging from 2.6 to 19.7, $p < .05$ in all cases).

Discussion of Analysis 3

Analysis 3 revealed that the pattern of performance shown by the five groups across 12 cognitive assessment measures may be explained with reference to three underlying factors: semantic impairment, a verbal production deficit, and a visual-processing impairment. These factors were confirmed by comparing the mean weightings from the different patient groups on each. Patients with known semantic deficits (SD) received lower weightings on Factor 1; those with known difficulties speaking (PNFA) had lower weightings on Factor 2; and those with known visual-processing impairments (PCA) received lower weightings on Factor 3.

Three further observations are of interest. First, the SD group did not differ from other patient groups on any but the first factor. It appears that, once the extent of semantic impairment is taken into account, little variability remains from the SD group to be accounted for by the other factors. In particular, SD did not—as PNFA did—receive low weightings on the verbal production factor; this outcome implies that, despite the large discrepancy between production and comprehension in SD from Analysis 2, impairments in SD can be explained with reference to a unitary underlying semantic impairment.

Second, the AD and fvFTD groups did not differ reliably from other patient groups on any of the three factors. For AD, it seems likely that this null result reflects the heterogeneous nature of the disorder: If AD affects each of the subsystems concurrently, then patients in this group may show mild to moderate impairments across the board and hence do not stand out in any individual factor. For fvFTD, their comparatively good performance on all tasks apart from fluency probably renders them difficult to differentiate on any of the three factors. There may well be other tasks not included in the current battery that are more exclusively influenced by frontal-executive dysfunction, which would more reliably discriminate the fvFTD group from the remaining patients.

Table 2
Rotated Component Matrix

	Factor 1	Factor 2	Factor 3
PPT-Word	.858		
WPM	.845		
VOSP-Sil	.826		.356
Naming	.824	.309	
PPT-Pics	.822		.323
CF	.748	.502	
DSF		.829	
LF	.523	.714	
MMSE	.558	.625	
VOSP-Inc. Let.		.428	.777
VOSP-Obj. Dec.	.349		.751
VOSP-Screen			.640

Note. Loadings less than 0.30 are suppressed. PPT-Words = Pyramids and Palm Trees test-Word Version; WPM = Word-Picture Match Test; VOSP-Sil = Silhouette naming from the Visual Object and Space Perception Battery; Naming = Naming test; PPT-Pics = Pyramids and Palm Trees Test-Picture version; CF: category fluency; DSF = toward digit span; LF = letter fluency; MMSE = Mini-Mental State Examination; VOSP-Inc. Let. = Incomplete letters test from the Visual Object and Space Perception Battery; VOSP-Obj. Dec. = Object Decision Test from the Visual Object and Space Perception Battery; VOSP-Screen = Screening Test from the Visual Object and Space Perception Battery.

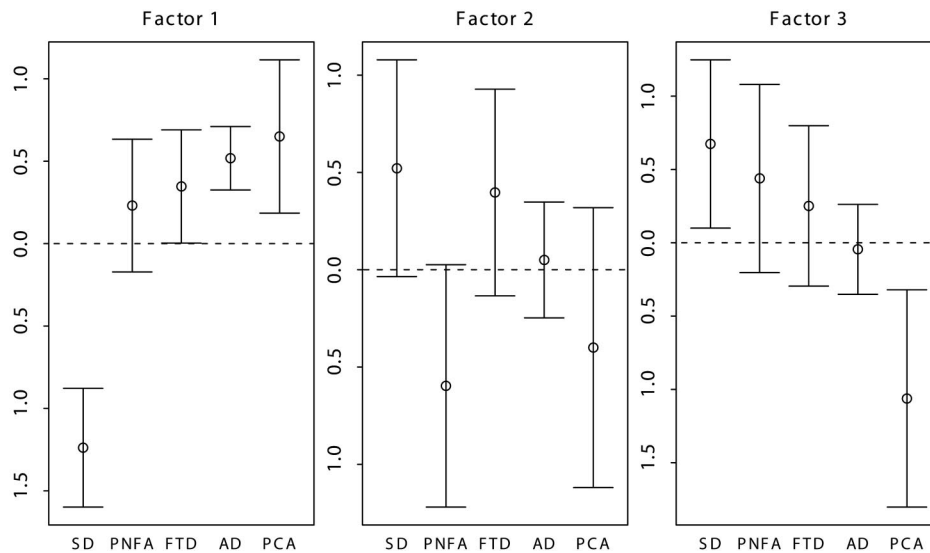


Figure 6. Confidence intervals of the mean factor loadings for each subject group on the three factors identified in Analysis 3. The first factor indexes semantic impairment; the second measures verbal production deficits; and the third taps visual-processing deficits. AD = Alzheimer's disease; CONT = controls; FTD = frontotemporal lobar degeneration; PCA = posterior cortical atrophy; PNFA = progressive nonfluent aphasia; SD = semantic dementia.

Third, it is worth pointing out that, across patient groups, semantic abilities are doubly dissociated from both visuospatial abilities and verbal production abilities. Specifically, comparison of SD and PNFA exemplifies a double-dissociation of semantic and verbal production abilities, whereas the contrast of SD and PCA exemplifies a double-dissociation of semantic and visuospatial abilities. These results again suggest that semantic impairment consists of something more than the conjoint impairment of visual perception and verbal production: Both of the latter faculties can be impaired in the context of comparatively good semantic knowledge and spared in the context of serious semantic memory impairment.

General Discussion

All five groups involved in the study showed deficits on the semantic battery, but important differences emerged that reflected the severity, purity, and consistency of the impairment (in terms of involvement of production and comprehension using verbal and visually based materials). Following is a summary of the patterns seen in the five syndromes in terms of these three key factors.

First, SD produces a moderate to severe semantic disorder that is observed across both visual and verbal modalities, for both production and comprehension. Patients with SD were reliably worse than other patient groups in all five semantic tasks and did not differ from the other groups in the nonsemantic tasks. Their great difficulty in the semantic tasks is not predominantly attributable to some combination of nonsemantic executive, verbal, and visual impairments. The evidence for this claim stems from the contrast of SD to three of the other patient groups. Unlike patients with known executive dysfunction (fvFTD), patients with SD showed substantial comprehension deficits and worse performance on category compared with LF. In contrast to patients with known

visuospatial deficits (PCA), the SD group did not reveal any discrepancy between visual and verbal tests of comprehension. Like the PNFA group, patients with SD showed a large discrepancy between comprehension and verbal production abilities; unlike this group, however, patients with SD do not suffer from a basic inability to produce any kind of fluent speech. The general pattern of impairment on receptive and expressive semantic tasks, coupled with relatively robust performance on nonsemantic tasks, suggests that the primary impairment in SD is to the semantic system. This conclusion also matches the results of the factor analysis, which demonstrated that, once the extent of their semantic impairment has been taken into account, patients with SD do not differ from other patient groups on other factors contributing to semantic task performance.

The findings suggest that the regions of predominant atrophy in SD—the anterior and inferolateral aspects of the temporal lobes (Chan et al., 2001; Davies, Graham, Xuereb, Williams, & Hodges, 2004; Galton et al., 2001a, b; Mummery et al., 2000)—must be a crucial component of the cortical semantic network, and they may encode the central conceptual knowledge that allows perceptual and language representations in different modalities of input and output to combine and communicate, as we have hypothesized elsewhere (e.g., Rogers et al., 2004). Therefore the profound expressive language deficit observed in SD is simply a reflection of the degradation of this anterior, bilateral temporal-lobe knowledge system (as previously suggested by Lambon Ralph et al., 2001), and it does not indicate an additional postsemantic deficit in word production.

Second, patients with PNFA have at most a mild semantic disorder that contrasts with their severe disorder of speech production. The current results suggest that the syndrome is relatively pure, but this may partly reflect the testing materials included in

the current battery. In contrast to SD, the neural basis of speech production deficits in PNFA is less clearly documented. Most studies have implicated left perisylvian structures (see Grossman, 2002); and a recent combined voxel-based morphometric MRI and FDG-PET study has shown consistent involvement of the anterior insula and inferior frontal region (Nestor, Graham, et al., 2003). This pathology is likely to impact on some additional cognitive abilities not adequately tested in the present study, such as executive function and working memory. With regard to the factors we have considered as supporting semantic task performance, impairments in PNFA appear to result primarily from an inability to produce fluent speech: Semantic tasks requiring verbal production were much more seriously impaired than were receptive semantic tasks; and even among patients with moderate or severe comprehension impairments, there was no indication that comprehension was more impaired for verbal than for visual materials.

Third, fvFTD has only a minor impact on the semantic system. The only tests for which the group was well down from control performance were category and LF, which, as we have noted, rely on the integrity and coordination of multiple cognitive faculties and were seriously impaired in all groups. Although fvFTD patients may be impaired at other kinds of executive tasks not included in the current battery, including the Stroop and Wisconsin Card Sorting Tasks (Perry & Hodges, 2000), the predominating clinical symptoms of the disorder involve social and personality abnormalities. The current results are generally consistent with a view of fvFTD as a disease that largely spares traditional cognitive functions, at least in mild and moderate phases. By contrast, AD is a syndrome involving significant impairment to almost every aspect of higher-level cognition, including episodic and semantic memory, attention and executive function, visual perception and praxis (Pasquier, 1999; Perry & Hodges, 2000; Welsh et al., 1992). Because the impact of the disease is so wide ranging and so heterogeneous in its manifestation across individuals (e.g., Martin et al., 1986), it is difficult to assess the integrity of any individual cognitive subsystem in this patient group. This is especially true for cognitive faculties like semantic memory, which may only be assessed by measurement tools that themselves presuppose the integrity of other cognitive resources. In the results we have described, the AD group was reliably impaired on all semantic and most nonsemantic tasks. The contribution of the current study, however, is to show that underlying semantic impairment can give rise to different patterns of impairment across tests, compared with performance deficits that result from nonsemantic factors. Specifically, patients with SD were comparatively worse on CF than LF, whereas this was not true of PNFA, fvFTD, or PCA groups. The AD pattern closely resembled the SD pattern, regardless of the overall severity of the comprehension deficit—suggesting that the great difficulty patients with AD encounter in verbal fluency arises partially or predominantly from an underlying semantic impairment, consistent with prior investigations (Monsch et al., 1992; Salmon et al., 1999). In general, the results agree with the notion that AD entails a consistent impairment to semantic memory, albeit somewhat milder than that observed in SD. When the severity of the comprehension impairment was taken into account, the AD group almost never differed from the SD group in the contrasts we investigated. Like patients with SD, they showed worse performance on CF than LF; equal degrees of impairment on visual and verbal tests of comprehension; and among cohorts

with mild or very severe comprehension deficits, equally impaired production and comprehension. The only contrast on which the groups differed substantially was among cohorts with moderate comprehension impairments: Patients with SD showed considerably poorer naming relative to their comprehension, whereas patients with AD did not—a finding that may be attributable to the predominantly left-lateralization of atrophy in SD, as discussed previously (Lambon Ralph et al., 2001).

Finally, the current results suggest that patients with PCA also have only very mild impairment in semantic memory, but they show striking problems with visually based tasks that are secondary to their visual-processing disorder. Apart from the clear impairment to visual perception revealed by poor performance on subtests of the VOSP, the PCA group differed strikingly from the typical AD group in two respects. First, patients with PCA showed equivalently poor performance in CF and LF, whereas patients with AD were considerably worse at CF than LF for all degrees of severity. Second, the PCA group showed a striking discrepancy between visual and verbal comprehension tasks, which increased with disease severity; the typical AD group showed equivalent performance in the two tasks. The results are consistent with the notion that PCA is a predominantly visuospatial disorder, and they reinforce the view that the cortical systems critical to semantic memory are the more anterior regions of the temporal lobes that are affected in SD and typical AD but not in PCA (at least in its initial phases; Mendez et al., 2002; Nestor, Caine, et al., 2003; Ross et al., 1996).

In conclusion, with regard to the stated aims of the study, the nature of semantic memory is such that any assessment of semantic abilities will always tax other cognitive abilities, such as visual processing (for pictorial stimuli), verbal processing (for spoken stimuli, responses, or both), and working memory or attentional processing (to deal with multiple response alternatives or the generation of serial responses). Poor performance on one or two semantic tasks does not provide a safe conviction of impaired semantic memory. The results we have described here suggest that, when performance is assessed across tasks that span visual and verbal modalities and tap mechanisms of reception and expression, the prospect of achieving an appropriate verdict is considerably brighter.

References

- Bayles, K. A., Tomoeda, C. K., Kaszniak, A. W., & Trosset, M. W. (1991). Alzheimer's disease effects on semantic memory: Loss of structure or impaired processing? *Journal of Cognitive Neuroscience*, 3(2), 166–182.
- Benson, D. F., Davis, R. J., & Snyder, B. D. (1988). Posterior cortical atrophy. *Archives of Neurology*, 45(7), 789–793.
- Bozdogan, H. (1987). Model selection and Akaike's information criterion: The general theory and its analytical extensions, *Psychometrika*, 52(3), 345–370.
- Bozeat, S., Gregory, C. A., Lambon Ralph, M. A., & Hodges, J. R. (2000). Which neuropsychiatric and behavioural features distinguish frontal and temporal variants of frontotemporal dementia from Alzheimer's disease? *Journal of Neurology, Neurosurgery and Psychiatry*, 69(2), 178–186.
- Braak, H., & Braak, E. (1991). Neuropathological staging of Alzheimer related changes. *Acta Neuropathologica*, 82(4), 239–259.
- Broe, M., Hodges, J. R., Schofield, E., Kersaitis, C., Shepherd, C. E., & Kril, J. J., et al. (2003). Staging disease severity in pathologically confirmed cases of frontotemporal dementia. *Neurology*, 60, 1005–1011.

- Caine, D., & Hodges, J. R. (2001). Heterogeneity of semantic and visuospatial deficits in early Alzheimer's disease. *Neuropsychology, 15*(2), 155–164.
- Caplan, D. (1987). *Neurolinguistics and linguistic aphasiology*. Cambridge, England: Cambridge University Press.
- Chan, A. S., Salmon, D. P., Butters, N., & Johnson, S. A. (1995). Semantic network abnormality predicts rate of cognitive decline in patients with probable Alzheimer's disease. *Journal of the International Neuropsychological Society, 1*, 297–303.
- Chan, D., Fox, N. C., Scahill, R. I., Crum, W. R., Whitwell, J. L., & Leschziner, G., et al. (2001). Patterns of temporal lobe atrophy in semantic dementia and Alzheimer's disease. *Annals of Neurology, 49*(4), 433–442.
- Chertkow, H., & Bub, D. (1990). Semantic memory loss in dementia of Alzheimer's type. What do various measures measure? *Brain, 113*(2), 397–417.
- Cockrell, J. R., & Folstein, M. F. (1988). Mini-Mental State Examination (MMSE). *Psychopharmacological Bulletin, 24*(4), 689–692.
- Coslett, H. B., Bowers, D., Verfaellie, M., & Heilman, K. M. (1991). Frontal verbal amnesia. Phonological amnesia. *Archives of Neurology, 48*, 949–955.
- Croot, K., Patterson, K., & Hodges, J. R. (1998). Single word production in nonfluent progressive aphasia. *Brain and Language, 61*(2), 226–273.
- Davies, R. R., Graham, K. S., Xuereb, J. H., Williams, G. B., & Hodges, J. R. (2004). The human perirhinal cortex and semantic memory impairment in neurodegenerative disease. *European Journal of Neuroscience, 20*(9), 2441–2446.
- Galton, C. J., Gomez Anson, B., Antoun, N., Scheltens, P., Patterson, K., & Graves, M., et al. (2001a). Temporal lobe rating scale: Application to Alzheimer's disease and frontotemporal dementia. *Journal of Neurology, Neurosurgery and Psychiatry, 70*(2), 165–173.
- Galton, C. J., Patterson, K., Graham, K., Lambon Ralph, M. A., Williams, G., & Antoun, N., et al. (2001b). Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. *Neurology, 57*(2), 216–225.
- Galton, C. J., Patterson, K., Xuereb, J. H., & Hodges, J. R. (2000). Atypical and typical presentations of Alzheimer's disease: A clinical, neuropsychological, neuroimaging and pathological study of 13 cases. *Brain, 123*(3), 484–498.
- Garrard, P., Lambon Ralph, M. A., Watson, P. C., Powis, J., Patterson, K., & Hodges, J. R. (2001). Longitudinal profiles of semantic impairment for living and nonliving concepts in dementia of Alzheimer's type. *Journal of Cognitive Neuroscience, 13*(7), 892–909.
- Garrard, P., Patterson, K., Watson, P. C., & Hodges, J. R. (1998). Category specific semantic loss in dementia of Alzheimer's type. Functional-anatomical correlations from cross sectional analyses. *Brain, 121*(4), 633–646.
- Glosser, G., & Friedman, R. B. (1991). Lexical but not semantic priming in Alzheimer's disease. *Psychology and Aging, 6*(4), 522–527.
- Graham, N. L., Patterson, K., & Hodges, J. R. (2004). When more yields less: Speaking and writing deficits in nonfluent progressive aphasia. *Neurocase, 10*(2), 141–155.
- Grossman, M. (2002). Frontotemporal dementia: A review. *Journal of the International Neuropsychological Society, 8*, 566–583.
- Hodges, J. R., Davies, R., Xuereb, J., Casey, B., Broe, M., & Bak, T., et al. (2004). Clinicopathological correlates in frontotemporal dementia. *Annals of Neurology, 56*, 399–406.
- Hodges, J. R., Davies, R., Xuereb, J., Kril, J., & Halliday, G. (2003). Survival in frontotemporal dementia. *Neurology, 61*, 349–354.
- Hodges, J. R., Graham, N., & Patterson, K. (1995). Charting the progression in semantic dementia: Implications for the organization of semantic memory. *Memory, 3*, 463–495.
- Hodges, J. R., & Miller, B. (2001). The neuropsychology of frontal variant frontotemporal dementia and semantic dementia. Introduction to the special topic papers: II. *Neurocase, 7*(2), 113–121.
- Hodges, J. R., & Patterson, K. (1995). Is semantic memory consistently impaired early in the course of Alzheimer's disease? Neuroanatomical and diagnostic implications. *Neuropsychologia, 33*(4), 441–459.
- Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia. Progressive fluent aphasia with temporal lobe atrophy. *Brain, 115*(6), 1783–1806.
- Hodges, J. R., Patterson, K., Ward, R., Garrard, P., Bak, T., & Perry, R., et al. (1999). The differentiation of semantic dementia and frontal lobe dementia (temporal and frontal variants of frontotemporal dementia) from early Alzheimer's disease: A comparative neuropsychological study. *Neuropsychology, 13*(1), 31–40.
- Hodges, J. R., Salmon, D. P., & Butters, N. (1992). Semantic memory impairment in Alzheimer's disease: Failure of access or degraded knowledge? *Neuropsychologia, 30*(4), 301–314.
- Howard, D., & Patterson, K. (1992). *Pyramids and Palm Trees*. Bury St. Edmunds, England: Thames Valley Test Company.
- Ikeda, M., Brown, J., Holland, A. J., Fukuhara, R., & Hodges, J. R. (2002). Changes in appetite, food preference, and eating habits in frontotemporal dementia and Alzheimer's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 73*(4), 371–376.
- Lambon Ralph, M. A., McClelland, J. L., Patterson, K., Galton, C. J., & Hodges, J. R. (2001). No right to speak? The relationship between object naming and semantic impairment: Neuropsychological evidence and a computational model. *Journal of Cognitive Neuroscience, 13*(3), 341–356.
- Lambon Ralph, M. A., Patterson, K., Graham, N. L., Dawson, K., & Hodges, J. R. (2003). Homogeneity and heterogeneity in mild cognitive impairment and Alzheimer's disease: A cross-sectional and longitudinal study of 55 cases. *Brain, 126*, 2350–2362.
- Martin, A. (1992). Semantic knowledge in patients with Alzheimer's disease: Evidence for degraded representations. In L. Backman (Ed.), *Memory functioning in dementia*. San Diego, CA: Elsevier Science Publishers, 119–134.
- Martin, A., Brouwers, P., Cox, C., & Fedio, P. (1985). On the nature of the verbal memory deficit in Alzheimer's disease. *Brain and Language, 25*, 323–341.
- Martin, A., Brouwers, P., Lalonde, F., Cox, C., Teleska, P., & Fedio, P., et al. (1986). Towards a behavioural typology of Alzheimer's patients. *Journal of Clinical and Experimental Neuropsychology, 8*(5), 594–610.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer's disease: Report of the NINCDS ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology, 34*(7), 939–944.
- Mendez, M. F., Ghajarania, M., & Perryman, K. M. (2002). Posterior cortical atrophy: Clinical characteristics and differences compared to Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders, 14*(1), 33–40.
- Mesulam, M. M. (1982). Slowly progressive aphasia without generalized dementia. *Annals of Neurology, 11*(6), 592–598.
- Mesulam, M. M., Grossman, M., Hillis, A., Kertesz, A., & Weintraub, S. (2003). The core and halo of primary progressive aphasia and semantic dementia. *Annals of Neurology, 54*(Suppl. 5), S11–S14.
- Monsch, A. U., Bondi, M. W., Butters, N., Salmon, D. P., Katzman, R., & Thal, L. J. (1992). Letter fluency exerts executive demands but depends less on semantic system: Comparisons of verbal fluency tasks in the detection of dementia of the Alzheimer type. *Archives of Neurology, 49*(12), 1253–1258.
- Mummary, C. J., Patterson, K., Price, C. J., Ashburner, J., Frackowiak, R. S., & Hodges, J. R. (2000). A voxel-based morphometry study of

- semantic dementia: Relationship between temporal lobe atrophy and semantic memory. *Annals of Neurology*, 47(1), 36–45.
- Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., & Black, S., et al. (1998). Frontotemporal lobar degeneration: A consensus on clinical diagnostic criteria. *Neurology*, 51(6), 1546–1554.
- Nebes, R. D. (1989). Semantic memory in Alzheimer's disease. *Psychological Bulletin*, 106(3), 377–394.
- Nebes, R. D., & Brady, C. B. (1990). Preserved organization of semantic attributes in Alzheimer's disease. *Psychology and Aging*, 5(4), 574–579.
- Nestor, P. J., Caine, D., Fryer, T. D., Clarke, J., & Hodges, J. R. (2003). The topography of metabolic deficits in posterior cortical atrophy (the visual variant of AD) with FDG-PET. *Journal of Neurology, Neurosurgery and Psychiatry*, 74(11), 1521–1529.
- Nestor, P. J., Fryer, T. D., Smielewski, P., & Hodges, J. R. (2003). Limbic hypometabolism in Alzheimer's disease and mild cognitive impairment. *Annals of Neurology*, 54, 343–351.
- Nestor, P. J., Graham, N. L., Fryer, T. D., Williams, G. B., Patterson, K., & Hodges, J. R. (2003). Progressive non-fluent aphasia is associated with hypometabolism centered on the left anterior insula. *Brain*, 126, 2406–2418.
- Ober, B. A., Shenaut, G. K., & Reed, B. R. (1995). Assessment of associative relations in Alzheimer's disease: Evidence for preservation of semantic memory. *Aging and Cognition*, 2(4), 254–267.
- Pasquier, J. (1999). Early diagnosis of dementia: Neuropsychology. *Journal of Neurology*, 246(1), 6–15.
- Perry, R. J., & Hodges, J. R. (1999). Attention and executive deficits in Alzheimer's disease. A critical review. *Brain*, 122(3), 383–404.
- Perry, R. J., & Hodges, J. R. (2000). Differentiating frontal and temporal variant frontotemporal dementia from Alzheimer's disease. *Neurology*, 54(12), 2277–2284.
- Plaut, D. C. (2002). Graded modality specific specialization in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology*, 19(7), 603–639.
- Rahman, S., Sahakian, B. J., Hodges, J. R., Rogers, R. D., & Robbins, T. W. (1999). Specific cognitive deficits in mild frontal variant frontotemporal dementia. *Brain*, 122(8), 1469–1493.
- Rogers, T. T., Lambson Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., & Hodges, J. R., et al. (2004). The structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, 111(1), 205–235.
- Rosen, H. J., Gorno-Tempini, M. L., Goldman, W. P., Perry, R. J., Schuff, N., & Weiner, M., et al. (2001). Patterns of brain atrophy in frontotemporal dementia and semantic dementia. *Neurology*, 58(2), 198–208.
- Ross, S. J., Graham, N., Stuart Green, L., Prins, M., Xuereb, J., & Patterson, K., et al. (1996). Progressive biparietal atrophy: An atypical presentation of Alzheimer's disease. *Journal of Neurology, Neurosurgery and Psychiatry*, 61(4), 388–395.
- Salmon, D. P., Heindel, W. C., & Lange, K. L. (1999). Differential decline in word generation from phonemic and semantic categories during the course of Alzheimer's disease: Implications for the integrity of semantic memory. *Journal of the International Neuropsychological Society*, 5(7), 692–703.
- Schwartz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 5(2), 461–464.
- Snowden, J. S., Goulding, P. J., & Neary, D. (1989). Semantic dementia: A form of circumscribed cerebral atrophy. *Behavioral Neurology*, 2, 167–182.
- Thompson, S. A., Patterson, K., & Hodges, J. R. (2003). Left/right asymmetry of atrophy in semantic dementia: Behavioural-cognitive implications. *Neurology*, 61, 1196–1203.
- Warrington, E. K., & James, M. (1991). *Visual Object and Space Perception Battery*. Bury St. Edmunds, England: Thames Valley Test Company.
- Welsh, K. A., Butters, N., Hughes, J. P., Mohs, R. C., & Heyman, A. (1992). Detection and staging of dementia in Alzheimer's disease. Use of the neuropsychological measures developed for the Consortium to Establish a Registry for Alzheimer's Disease. *Archives of Neurology*, 49(5), 448–452.

Received April 19, 2004

Revision received April 28, 2005

Accepted May 2, 2005 ■