doi:10.1093/brain/awl069 Brain (2006), **129**, 1152–1163

Lost in semantic space: a multi-modal, non-verbal assessment of feature knowledge in semantic dementia

Peter Garrard[1,2] and Erin Carroll[1]

[1]University College London, Institute of Cognitive Neuroscience and [2]Royal Free Hospital, Hampstead, London, UK

Correspondence to: Dr Peter Garrard, Institute of Cognitive Neuroscience, Alexandra House, 17 Queen Square, London WC1N 3AR, UK
E-mail: p.garrard@ucl.ac.uk

A novel, non-verbal test of semantic feature knowledge is introduced, enabling subordinate knowledge of four important concept attributes—colour, sound, environmental context and motion—to be individually probed. This methodology provides more specific information than existing non-verbal semantic tests about the status of attribute knowledge relating to individual concept representations. Performance on this test of a group of 12 patients with semantic dementia (10 male, mean age: 64.4 years) correlated strongly with their scores on more conventional tests of semantic memory, such as naming and word-to-picture matching. The test's overlapping structure, in which individual concepts were probed in two, three or all four modalities, provided evidence of performance consistency on individual items between feature conditions. Group and individual analyses revealed little evidence for differential performance across the four feature conditions, though sound and colour correlated most strongly, and motion least strongly, with other semantic tasks, and patients were less accurate on the motion features of living than non-living concepts (with no such conceptual domain differences in the other conditions). The results are discussed in the context of their implications for the place of semantic dementia within the classification of progressive aphasic syndromes, and for contemporary models of semantic representation and organization.

Keywords: connectionist modelling; motion; semantic dementia; semantic features; semantic memory

Abbreviations: FRT = Feature Reality Test; CCP = Camel and Cactus Test

Original Submission July 19, 2005. Revised December 16, 2005. Accepted March 1, 2006. Advance Access publication April 3, 2006

Introduction

It is widely agreed that the syndrome of semantic dementia is most usefully classified within the group of progressive, non-Alzheimer neurodegenerative conditions in which language is the major focus of functional deterioration (Hodges *et al.*, 1992; Hodges and Patterson, 1996; Kertesz *et al.*, 2003; Grossman and Ash, 2004). Individuals with semantic dementia typically present either complaining of word-finding difficulties ('forgetting words'), and/or with less specific communication problems endorsed by a spouse or other relation. In the early stages, spontaneous discourse is often strikingly well preserved: phonologically and grammatically correct sentences are produced at a normal rate, though usually with an excess of high frequency, generic words or phrases substituting for specific content-bearing words (e.g. 'the stuff you have' or 'those things in there').

Neuropsychological evaluation reveals a highly characteristic pattern of deficits incorporating (i) severe anomia and impaired single-word comprehension; (ii) preservation of other aspects of linguistic processing such as phonology and syntax; and (iii) relative sparing of day-to-day memory, non-verbal problem-solving and visuospatial abilities (Snowden *et al.*, 1989, 1996; Hodges *et al.*, 1992). Together with a highly typical pattern of asymmetric (usually left-sided) atrophy of the anterior and inferior aspects of the temporal lobe, these clinical and neuropsychological features allow semantic dementia to be accurately differentiated from Alzheimer's disease during life in the majority of cases (Galton *et al.*, 2000; Hodges *et al.*, 2004; Rosen *et al.*, 2005).

Experimentally motivated investigation of this fascinating group of patients has resulted in a detailed understanding of the deficit at the cognitive level, as well as its impact on other language tasks. For instance, it is clear that (i) the patient's responses on individual test items do not suggest that conceptual representations are either 'present or absent', but rather show greater or lesser degrees of partial degradation (Warrington, 1975; Hodges *et al.*, 1995); and (ii) that, far from impacting exclusively on the appreciation of word meaning, semantic dementia also interferes with the production or comprehension of stimuli in non-verbal task domains, such as drawing, picture–picture association, and colour-object or colour-sound matching (Breedin *et al.*, 1994; Bozeat *et al.*, 2000, 2003; Lambon Ralph and Howard, 2000).

It has also been argued that the regularization errors typically made by semantic dementia patients when required to transform linguistic information between representational domains (e.g. surface dyslexic/dysgraphic errors in reading/writing, and difficulty with irregular morphosyntactic derivatives) have their origin in the reduced influence of degraded semantics within highly interactive mapping systems (Patterson and Hodges, 1992; Patterson *et al.*, 2001). More recently, Rogers *et al.* (2004*b*) demonstrated that the ability to distinguish between words and non-words, and between real and chimeric objects is also modulated by semantic impairment, particularly for low-frequency items with atypical orthographic or physical features.

Rogers *et al.* (2004*a*) argued that the range of phenomena just described could be viewed as arising from the degradation of an amodal system whose function is to allow representations in multiple domains, each subject to different degrees of regularity, reliably to map one with another. This overarching idea was supported by the behaviour of a connectionist model that incorporated its theoretical assumptions in a recurrent architecture. Briefly, semantic information was represented in the weights learned by the model's hidden units (see Fig. 1), which, in the absence of direct external inputs from the training environment (partially overlapping sets of visual or verbal features), were sensitive to second-order similarity. Hidden units thus came to acquire '...abstract representations, whose similarity relations are not tied to any individual modality, but capture the deep structure across modalities' (Rogers *et al.*, 2004*a*, p. 206).

After increasing degrees of disruption to the weights learned by the hidden layer during the training process, the model exhibited patterns of performance that were strikingly similar to those seen in patients with semantic dementia. Specifically, the tasks of mapping between layers (i.e. reinstating a learned association between an activation pattern in one set of units and an activation pattern in another) resulted in error rates that were related to the degree of hidden weight disruption in a fashion that was qualitatively and quantitatively similar to longitudinal patterns of decline documented for semantic dementia.

The authors argued that the patterns of performance in patients and in the degraded model resulted from a common underlying factor—to wit, the dynamics of processing in a distributed system based on higher-order similarity structure. For reasons of computational economy only two modalities

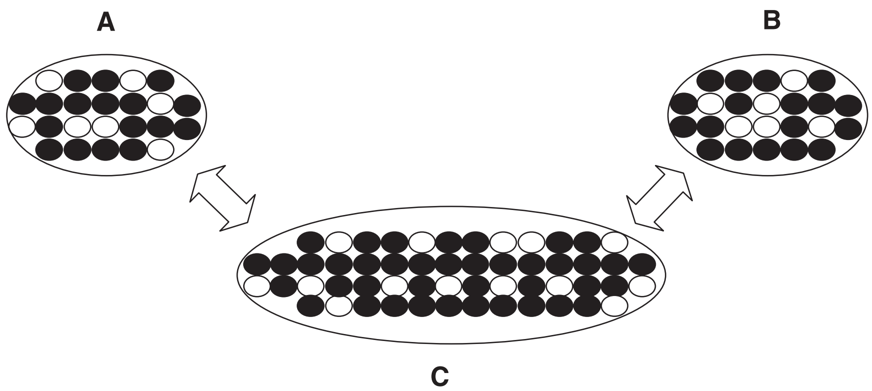

Fig. 1 Diagrammatic representation of the architecture of the recurrent connectionist model described by Rogers *et al.* (2004). **A**, **B** and **C** represent groups of individual units with bidirectional weighted connections, in either active (open oval circle) or inactive (filled oval circle) states. Different patterns of activation over units in **A** and **B** stand for the representation in some sensory modality of individual concepts, such that their real-world similarity is reflected in the overlap of active units in the model. For instance, a visual representation designated 'horse' would have many active units in common with the representation for 'donkey', fewer with the representation for 'elephant' and fewer still with that assigned to 'snake'. By systematically adjusting the weights between these peripheral units and the intermediate—or 'hidden'—units (**C**), the model is trained such that units corresponding to the same concept at **A** and **B** become mutually activating. The 'knowledge' that allows the activation pattern associated with 'horse' to appear at **B** when the appropriate pattern is presented to **A** (and vice versa) is thus inherent in the pattern of learned weights in the connections between **A** and **C**, and **B** and **C**.

(verbal and visual) were incorporated into the model, allowing the tasks of mapping between these two information sources (i.e. the analogues of picture naming, and word-to-picture matching) to be simulated. Its theoretical predictions, however, can be extended to include other forms of information (auditory, tactile, etc.). Indeed, inaccuracy in achieving many of these specific 'mappings' have already been documented in semantically impaired individuals (Luzzatti and Davidoff, 1994; Bozeat *et al.*, 2000; Lambon Ralph and Howard, 2000; Miceli *et al.*, 2001; Coccia *et al.*, 2004).

It is also worth noting that the anterior temporal focus of pathology in semantic dementia is consistent with the notion of damage to a supramodal system: the temporal pole has connections with inferior, middle and superior temporal gyri, which act, respectively, as the terminus for the ventral visual processing stream, an integrative region for somatosensory, visual and auditory processing streams, and a centre of high-level auditory processing (Gloor, 1997).

Compelling as this theoretical model may be on behavioural, computational and neuroanatomical grounds, it has recently been suggested that the syndrome of semantic dementia may be more heterogeneous, representing in some cases a verbal comprehension deficit combined with a degree of visual agnosia, and in others a form of progressive degeneration in neural systems governing language output associated with a fluent clinical phenotype (Mesulam, 2001).

For a variety of reasons, this claim has proved difficult to refute on empirical grounds. A purely verbal method of probing a semantic dementia patient's knowledge of a concept's semantic attributes [as employed, for instance, by Garrard *et al.* (2005) and Hodges *et al.* (1996) to explore the extent of semantic impairment in Alzheimer's disease] would be limited by the profound degree of anomia exhibited by all but the most mildly impaired subjects: does a failure to endorse the statement 'an elephant has a trunk' stem from a degraded representation of the conceptual representation of an elephant, or from a failure to comprehend the nominal term 'trunk'? Failure on purely visual analogues of the same task [e.g. the Pyramids and Palm trees (Lambon Ralph *et al.*, 2001) or the Camel and Cactus (Bozeat *et al.*, 2000)] may, likewise, be due to agnosic difficulties with the representation either of the subject or of its predicate. Tasks requiring mapping between specific domains, such as those mentioned above, have documented consistent impairment across pairs of individual feature domains, usually in single cases, but do not address the issues of consistency between cases or generalization across domains.

The study presented here was intended to overcome these objections as far as possible, by (i) probing a common set of items across more than one non-verbal modality in a group of patients with semantic dementia; and (ii) correlating performance on individual items with the patients' ability on the paradigmatic verbal task of picture naming. In adopting this approach, we set out to achieve evidence of consistency, both within individual patients across different modalities, and across the patient population as a whole, and thus to provide empirical support for the theoretical model of semantic dementia as a homogeneous manifestation of disruption to a unitary system of supramodal representations; in short, to show that the syndrome of semantic dementia emerges as the patient becomes progressively lost in a multi-dimensional 'semantic space'.

Participants, materials and methods

Patients

Twelve right-handed patients meeting clinical and radiological criteria for semantic dementia (Hodges *et al.*, 1992) were recruited from specialist clinics in London and Cambridge. Ten were male. Ages ranged from 56 to 77 years, with a mean of 64.4 (6.8 SD). Patients had completed a mean of 13 years of full-time education (3.2 SD). Informed consent was obtained from each patient or, where appropriate, from their carer. The study was approved by local research ethics committees in London and Cambridge.

Controls

Twenty-three control participants (two left-handed) were selected from a volunteer database to match as closely as possible the age and education level of the patient group. The male to female ratio was 10 : 13. The mean age was 64.7 (5.1 SD) and mean years of education 13.6 (3.3 SD). Exclusion criteria consisted of an MMSE score of <26, and a concurrent or previous history of head injury or stroke, major neurological or psychiatric illness, or alcohol abuse.

Background neuropsychological tests

Patients were assessed using the semantic battery described by Bozeat *et al.* (2000). This battery employs a set of 64 common, concrete concepts drawn from the Snodgrass corpus (Snodgrass and Vanderwart, 1980), which can be presented in either picture or word form for naming, word-to-picture matching, sorting and associative matching. The battery contains equal subsets of living and non-living items, matched for familiarity and age of acquisition, allowing category-specific deficits to be identified.

Visuospatial function and colour knowledge were assessed using subtests of the Visual Object and Space Perception battery (VOSP) (Warrington and James, 1991), and the colour knowledge test of De Vreese *et al.* (1994).

Experimental task—Feature Reality Test (FRT)

Subjects were presented with pairs of pictures, and asked to indicate the 'more real looking' of the two. The two pictures in each presentation differed along a single featural dimension. Thirty-three concrete, picturable concepts, from animate and inanimate domains, were selected such that all appeared in at least two out of four test conditions (see Table 1).

Successful performance in each condition required knowledge of a specific stimulus feature, as follows:

Condition 1: Colour

Twenty-nine concepts were presented in the form of coloured drawings. In each test item, an appropriately coloured item was paired with the same item pictured in an inappropriate colour. Each target item was presented twice—once together with a distractor in a colour that was attributable to other semantically related items

Table 1 Items appearing in the FRT, associated familiarity ratings, and the conditions in which each was presented

Item	Semantic domain	Mean familiarity rating*	Colour	Context	Sound	Motion	Semantic battery
Apple	i	4.90	■	■	■		■
Banana	i	4.48	■	■	■	■	■
Bee	a	2.43	■		■		
Cat	a	3.62	■	■	■	■	■
Clock	i	4.86			■	■	
Cow	a	2.29	■	■	■	■	■
Crab	a	1.38	■	■		■	
Dog	a	3.52	■	■	■	■	■
Duck	a	2.48	■	■	■	■	■
Egg	i	4.47	■	■	■	■	
Elephant	a	1.43	■	■	■		■
Fire engine	i	2.86	■	■	■		
Frog	a	2.14	■	■	■	■	■
Guitar	i	2.23	■	■	■		
Hammer	i	2.90		■	■	■	■
Helicopter	i	2.57	■	■	■	■	■
Horse	a	2.86	■	■	■	■	■
Kangaroo	a	1.38	■	■	■	■	■
Mouse	a	1.90	■	■	■		■
Orange	i	4.00	■	■			■
Penguin	a	1.38	■	■	■	■	■
Piano	i	3.33	■	■		■	■
Pig	a	1.86	■	■	■		
Pigeon	a	3.76	■	■			
Saw	i	2.33		■	■	■	■
Scissors	i	4.48	■		■	■	■
Snowman	i	2.10	■	■			
Suitcase	i	3.76	■	■	■	■	■
Tap	i	4.95			■		
Tiger	a	1.33	■	■	■		■
Tortoise	a	1.38	■	■		■	■
Trumpet	i	1.67	■	■	■		
Windmill	i	1.71	■	■		■	
Mean (SD) familiarity	Animate	2.20	2.20	2.18	2.20	2.24	2.14
		(0.86)	(0.86)	(0.88)	(0.78)	(0.88)	(0.83)
	Inanimate	3.39	3.27	3.09	3.57	3.49	3.64
		(1.17)	(1.15)	(1.08)	(1.18)	(1.09)	(0.91)
	P-value	<0.01	<0.01	<0.05	<0.01	<0.05	<0.01

Filled cells indicate that the item was represented in the corresponding condition; a = animate; i = inanimate. *Ratings were obtained from the control group using the method described by Snodgrass and Vanderwart (1980).

('plausible condition') and once with a distractor in a colour that did not meet this criterion ('implausible condition'—plausible and implausible foils from the colour and context conditions were compared using rated typicality scores provided by 20 of the control subjects. As expected, plausible items were associated with significantly higher scores on this measure).

For example, a yellow banana was shown with one coloured orange (plausible) or one coloured pink (implausible).

Condition 2: Environmental context

Twenty-nine concepts were presented in the form of photographs superimposed on an environmental context. Target items were shown in a typical context, and distractors in both plausibly and implausibly incorrect contexts (defined as for Condition 1, above). For example, a horse was shown superimposed on a field, on a desert scene (plausible), or in a shopping centre (implausible).

Condition 3: Sound

Twenty-six concepts were presented in the form of pairs of black and white line drawings accompanied by a recorded sound. Target concepts were again paired with plausible and implausible distractors—the former from the same semantic domain (living versus non-living) and the latter from the opposite domain. For example, the sound of a dog barking would accompany a picture of a dog, a cat (plausible) or an alarm clock (implausible).

Condition 4: Motion

Twenty concepts were presented in the form of pairs of animations, each showing the same item involved in two distinct kinds of motion. For example, a cow was depicted as chewing and swinging its tail from side to side in the target animation, but as rearing up on its hind legs in the manner of a horse, in the distractor image. The plausibility manipulation was not incorporated into the motion test.

Table 2 Performance on general neuropsychological assessments and semantic battery subtests

Test	Control mean (SD)‡	Subject											
		BG	JM	AN	NG	VH	DW	RJW	WM	TW	FO	IB	AT
MMSE (Max 30; Normal > 26)		27	28	27	n/a	27	25	24	21	23	n/a	23	15
GNT (Max 30)		6*	3*	1†	n/a	11*	0†	0†	0†	0†	0†	0†	0†
VOSP													
Screening test (Max 20)		20	19	20	19	20	18	18	20	n/t	n/t	n/t	19
Incomplete letters (Max 20)		16	19	20	20	20	20	n/t	18	n/t	12	n/t	18
Object decision (Max 20)		n/t	n/t	n/t	19	n/t	14	17	n/t	n/t	n/t	n/t	n/t
Dot counting (Max 10)		8	9	10	10	10	10	9	10	10	n/t	10	10
Position discrimination (Max 20)		19	20	20	10	20	19	20	20	n/t	n/t	n/t	n/t
Number location (Max 20)		6*	10	10	10	10	10	6*	10	n/t	n/t	n/t	n/t
Cube analysis (Max 20)		7	9	10	9	10	10	0*	10	10	10	10	10
Colour Battery (De Vreese *et al*., 1994)													
Colour naming /10		8	10	n/t	10	10	9	7	n/t	n/t	n/t	7	n/t
Colour identification /10		10	10	n/t	10	10	10	10	n/t	n/t	n/t	9	n/t
Picture naming													
• Living (Max 32)	62.3	19	20	19	31	17	10	23	5	4	1	2	0
• Non-living (Max 32)	(1.6)	22	28	22	29	21	23	16	5	4	0	3	0
Word-to-picture matching													
• Living (Max 32)	63.7	30	30	31	27	27	20	28	4	12	5	3	n/t
• Non-living (Max 32)	(0.5)	31	31	31	32	32	29	28	14	26	6	16	
• CCP (Max 64)	56.9 (7.3)	55	54	49	46	41	35	35	33	27	23	19	n/t

All scores refer to a period within six months of administration of the experimental battery. MMSE = Mini Mental State Examination (Folstein *et al*., 1983); GNT = Graded Naming Test (McKenna and Warrington, 1983); VOSP = Visual Object and Space Perception battery (Warrington and James, 1991); CCP = Camel and Cactus test (picture version) (Bozeat *et al*., 2000); n/t = not tested; n/a = result not available. *Performance below 5th percentile for age group. †Performance below 1st percentile for age group. ‡Control values refer to performance on both living and non-living subsets—taken from Bozeat *et al.* (2000).

The test was administered in blocked format, one feature condition at a time, with the order of trials randomized within each. Stimuli were presented, horizontally separated, on a laptop computer screen. Subjects were instructed to indicate their responses using the left or right arrow keys on the computer's keyboard.

Results

Standard semantic tasks

Table 2 displays the scores achieved by the patients on a general neuropsychology assessment and subtests from the semantic battery, together with norms from age- and education-matched controls. The cases are arranged in order of increasing severity, as indexed by the picture version of the Camel and Cactus test (CCP). As the table indicates, patients exhibited a range of severity of semantic impairment, and several showed poorer performance with items from the living subset (DW, IB, TW, WM, JM) on at least one of the subtests. These findings are consistent with those reported by Lambon Ralph *et al.* (2003) from a similar patient group.

Reality decision

Controls averaged 95.71% (2.47 SD) correct in the colour condition of the FRT, 96.52% (2.80 SD) in the environmental context condition, 94.57% (2.94 SD) in the sound condition and 97.93% (3.26 SD) in the motion condition. The patients' scores averaged: colour—81.5% (12.3 SD); environmental context—81.56% (11.8 SD); sound—77.27% (13.1 SD) and motion—77.46% (13.2 SD).

A 2 (group) by 4 (condition) ANOVA (analysis of variance) (with data from case FO excluded because of chance-level performance in all four test conditions) confirmed that the overall difference between groups was significant [$F(1,33) = 56.6$, $P < 0.001$], but that the difference across conditions [$F(3,33) = 1.4$, $P = 0.25$] and the group by condition interaction [$F(3,33) = 1.6$, $P = 0.2$] both fell short of significance. Each patient's scores on the four feature conditions are shown, after conversion to z-scores, in Fig. 2. All but the most mildly affected case (NG, who was able to name 60 out of the 64 items in the semantic battery) fell two or more standard deviations below the control average on at least one of the test conditions. With the apparent exception of cases AN and WM, whose performance was strikingly worse on motion than on the other three conditions, the individual patterns of performance across conditions were generally homogeneous.

Rogers *et al.* (2003*a*) showed that patients with semantic impairments can perform normally on an object decision task if the target items are structurally prototypical and distractors structurally unusual. This important observation was incorporated into the colour, context and sound conditions of the FRT by manipulating the plausibility of the distractor item (see Methods). The effect of distractor plausibility on patient and control responses was examined in a separate

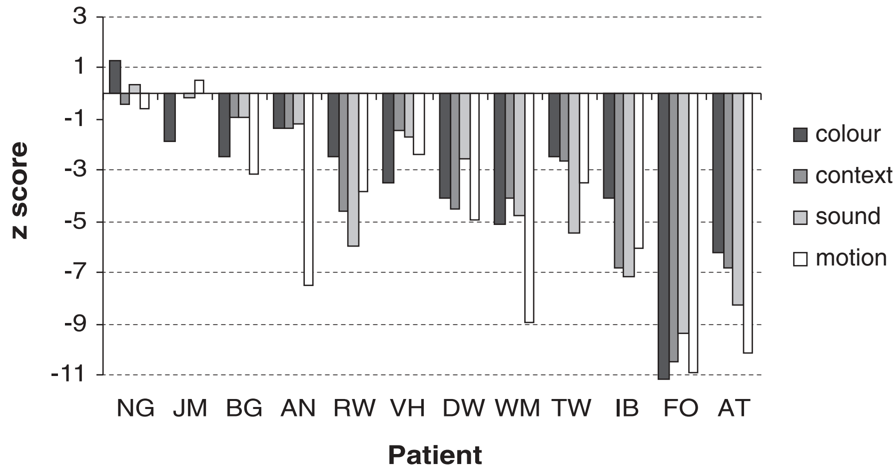

Fig. 2 Performance on the four conditions of the FRT by 12 patients with semantic dementia. Values are represented as z-scores based on the scores achieved by a group of age-matched, cognitively normal controls. Cases are arranged from left to right in terms of decreasing performance on the naming test of the semantic battery.

repeated-measures ANOVA with plausibility as a two-level, within-subjects factor. In accordance with the findings of Rogers *et al.* (2003*a*), plausible distractors were correctly rejected less often than implausible [$F(1,32) = 119.877$, $P < 0.01$], a difference that was exaggerated in the patient group, resulting in a significant group by plausibility interaction [$F(1,32) = 41.927$, $P < 0.01$].

Comment

Although a correct response may be achieved purely by chance, an erroneous response in any of the FRT conditions indicates a failure to map successfully between the visual form and the modality of the attribute represented. In the context of semantic dementia, it is assumed that this failure of mapping is due to the operation of a degraded semantic system, rather than to deficits referable to the input systems themselves. Three predictions follow from this assumption: first, performance on each item should demonstrate a high degree of consistency between conditions (i.e. an incorrect decision in one condition predicts failure in all); secondly, patients' performance on the FRT should correlate with their scores on standard semantic tests such as naming and word-picture matching; and finally, the number of correctly attributed features should be expected to be higher with items for which semantic knowledge is sufficiently preserved to support a correct naming response than with those that the patient is unable to name.

Further analyses were performed on the data to test each of these predictions in turn.

Analysis 1: Consistency between conditions

Individual likelihood ratio χ^2 coefficients were calculated for each patient's responses (correct or incorrect) on the four feature conditions. Values of $\chi^2(3)$ ranged from 1.39 (case VH) to 6.14 (case AN), but all fell short of statistical significance, indicating similar levels of performance across conditions. A more detailed item-by-item analysis, based on pooled responses from the entire patient group was conducted to establish the level of response consistency across conditions for items appearing in all four, three out of four and two out of four feature conditions (see Table 1).

Although the value of χ^2 from the first of these three comparisons did not differ significantly from a random distribution of errors [$\chi^2(3) = 2.96$, $P > 0.05$], significant consistency across conditions was demonstrated for the items that appeared in three feature conditions [$\chi^2(2) = 6.59$, $P < 0.05$], and in two [$\chi^2(1) = 19.32$, $P < 0.01$].

To assess the response consistency across conditions in individual cases, the responses of each patient were examined, item-by-item, using logistic regression to identify cases where feature type was independently predictive of success. Feature type, conceptual domain (living versus non-living) and concept familiarity ratings were entered into the model. The analysis revealed only one patient (AT)—whose test performances consistently fell in the severely impaired range—in whom feature-reality condition was predictive of a correct response (Wald = 8.771, $P < 0.05$). Analysis of AT's scores also revealed the influence of the interaction between feature and domain on his responses (Wald = 8.049, $P < 0.05$) due to a smaller number of correct responses on the inanimate than animate test item concepts in the motion condition.

Comment

The cross-modal consistency shown in the above analyses are assumed to support the degradation of a supramodal system as the functional basis of semantic dementia. It is still possible, however, to construe the results as reflecting an artefact of concept familiarity—a variable well known to influence performance on semantic tasks (Funnell and Sheridan, 1992). This is because even if semantic dementia did reflect the parallel degradation of multiple functionally independent, domain-specific subsystems, it would still be the highly familiar items that would tend to be preserved in each, producing a similar item-wise correlation in individual cases.

In a further analysis, item-wise correspondence was sought between performance on the FRT and picture naming. To examine the 'familiarity hypothesis', this procedure was carried out in two ways: first across all available test items, and secondly on a restricted subset of item pairs matched for familiarity. Under the 'familiarity hypothesis', FRT scores will be higher for items that could be named for no other reason than that all aspects of the named items are generally more familiar; this difference should, therefore, fail to emerge when the named and unnamed items in question are of comparable (high or low) familiarity. In contrast, if cross-modal consistency is a consequence of supramodal disruption, then the association between named items and higher scores on the FRT should persist, even after familiarity matching.

Analysis 2: Correspondence between naming and feature knowledge

In two separate studies of patients with Alzheimer's disease, Hodges *et al.* (1996), and more recently Garrard *et al.* (2005), reported a strong correspondence between a patient's ability to name a concept from its picture, and the quality and quantity of information that he/she could produce in response to its spoken name. This result was interpreted as providing support for the notion that anomia in Alzheimer's disease is underpinned, at least in part, by a degradation in semantic knowledge about common concepts. The profound anomia typically evinced by patients with semantic dementia even at relatively early stages would make such a finding difficult to reproduce using these strictly verbal techniques, yet the relationship would be expected to hold at least as strongly. The extensive overlap between items probed in the FRT and those that form the basis of the semantic battery meant that a similar relationship could be sought between naming and this non-verbal measure of feature knowledge.

Twenty-one of the thirty-three concepts used in the FRT had also appeared in the picture naming subtest of the semantic battery (see Table 1). This stimulus overlap provided an opportunity to test the hypothesis that performance on the FRT would be better for named than unnamed items. For this comparison, an aggregated feature reality score, based on the proportion of correct responses given, was assigned to each concept. Collapsing across patients resulted in 141 correct and 111 incorrect naming responses. As shown in Fig. 3A, a significantly higher proportion of correct feature reality responses were recorded on named than unnamed items [$t(250) = 3.91$, $P < 0.001$]. The difference remained significant in a by-subjects analysis, in which each patient's FRT scores on named and unnamed items were subjected to a paired *t*-test [$t(20) = 4.1$, $P < 0.01$].

The relationship between naming and feature knowledge was also examined condition by condition. This was achieved by categorizing each response as correct if the target was chosen in preference to both the plausible and the implausible distractors, and incorrect if either of the two distractors was chosen in preference to the target. The proportions of correct responses were greater with named than unnamed items in all conditions (Fig. 3B), though comparisons were statistically significant only in the colour [$\chi^2(1) = 9.45$, $P < 0.01$] and sound [$\chi^2(1) = 9.62$, $P < 0.01$] conditions.

The values of rated familiarity for the semantic battery item subset ranged from 1.3 [tiger] to 4.9 [apple], with a median value of 2.6, a mean of 2.8 and a standard deviation of 1.1, allowing the predictions of the 'familiarity hypothesis'

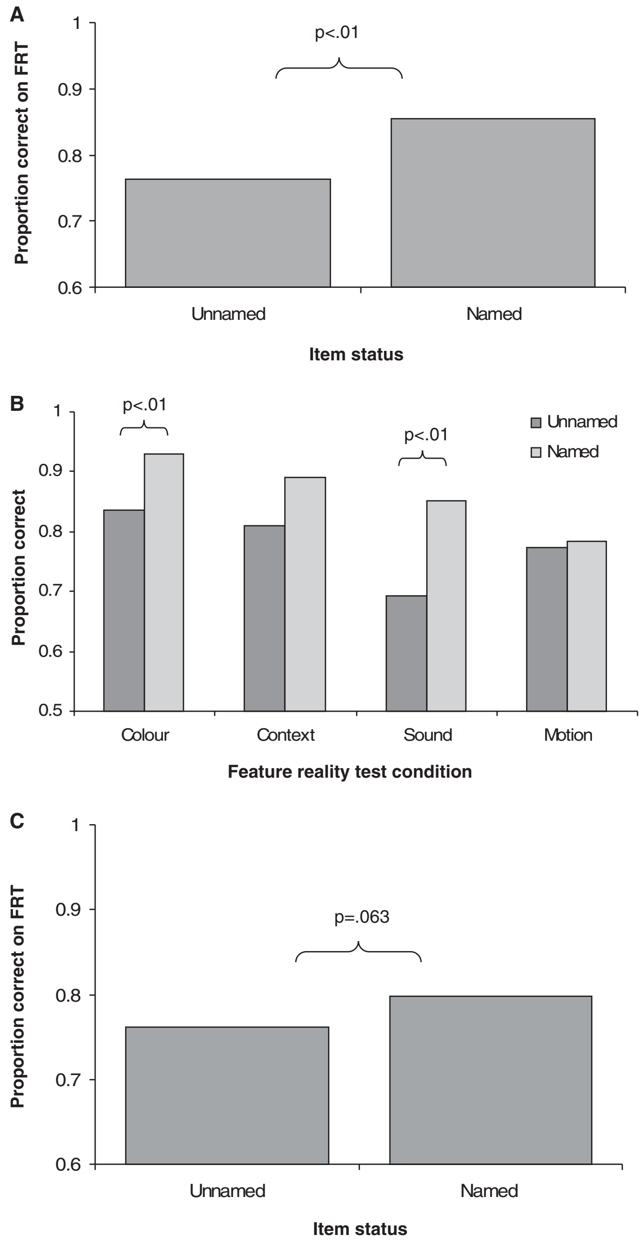

Fig. 3 Performance on the FRT for those items that could be named compared with those that could not. Proportional performances are shown for the test as a whole (**A**), for scores achieved in each of the four conditions of the FRT (**B**) and using only data from familiarity-matched item pairs (**C**).

discussed above to be tested by selecting pairs of items with similar familiarity ratings among each patient's responses, such that the patient had correctly named one item but failed to name the other. This was accomplished by pairing each of the 21 test items in each patient's set of naming responses with every other item that met the following criteria: (i) having a familiarity rating within 0.5 standard deviations of the index item; and (ii) associated with an outcome on naming that was opposite to that of the index item. These criteria were met by 152 item pairs, drawn from the responses of 9 out of the 11 patients (no pairs were formed from the responses of NG, who named all the items in the set, or from those of AT, who named none).

As shown in Fig. 3C, the higher FRT scores associated with named items persisted, though the difference was smaller and fell marginally short of statistical significance using a paired *t*-test [$t(151) = 1.87$, $P = 0.063$]. The direction of the trend would, however, favour the conclusion that the difference between named and unnamed items is not entirely due to the effect of familiarity.

Analysis 3: Correlation between feature reality and other semantic tests

Percentage scores achieved by ten of the patients on each of the feature conditions were entered into a correlation analysis with scores on the naming, word-to-picture matching and Camel and Cactus (picture) subtests from the semantic battery. Data from AT and FO were omitted, as the former did not complete all the semantic battery tasks, and the latter performed at chance level on the FRT. Separate analyses were conducted using responses on trials with plausible and implausible foils. The results (displayed in Table 3) confirmed that overall performance on the FRT was highly correlated with all three of the conventional measures, and that correlations were considerably stronger for items with plausible distractors [an observation also made by Rogers *et al.* (2004*b*) in connection with performance on object decision].

When the four test conditions were considered separately, the pattern of correlations persisted with the colour, context and sound conditions (with the exception of the correlation between colour and CCP, which fell just short of significance). In contrast, performance on the motion condition did not correlate significantly with any of the conventional measures, suggesting that this domain, or the stimuli through which it was probed, may be different from the others. This difference is explored in more detail below.

Analysis 4: Interactions between feature modality and semantic domain

The results of Analysis 3 included the finding that performance on the motion reality test did not correlate with any of the more conventional measures, while significant correlations were obtained between the remaining three feature conditions and almost all semantic battery tasks. The motion condition of the FRT yielded similarly anomalous results in the comparisons between named and unnamed items: motion was no more likely to be correctly attributed to a concept if the patient had named that item than if he/she had failed to do so.

One possible reason for these differences is that the importance of individual features varies between concepts of different types. This assumption is central to a number of existing theories, including that of Rogers *et al.* (2004*a*). While making no specific claims about the issue of category-specific semantic impairment, Rogers *et al.* did point out that living and non-living concepts are likely to differ in terms of the amount of structure they share in various representational domains. In the representation of action, for instance, there is likely to be more shared structure for artefacts than living things: 'Objects that afford similar actions may induce similar representations in areas of cortex that subserve action, and this structure may also constrain the similarity relations acquired by the semantic system as it learns the mappings between object appearances and appropriate actions' (p. 231).

It is conceivable that the type of motion associated with an item forms a further example of a representational domain in which such asymmetry of shared structure has consequences for similarity at the semantic level. Intuitively, motion might be considered as more 'causally relevant' to the living than the

Table 3 Pearson correlations between each of the features in the FRT and subtests from the semantic battery

	FRT (plausible)	FRT (implausible)	FRT (colour)	FRT (context)	FRT (sound)	FRT (motion)	Naming	WPM	CCP
FRT (all conditions)	0.929**	0.744*	0.850**	0.873**	0.782**	0.487	0.847**	0.784**	0.742*
FRT (plausible)		0.741*	0.801**	0.878**	0.915**	0.207	0.934**	0.891**	0.853**
FRT (implausible)			0.560	0.949**	0.818**	−0.067	0.532	0.537	0.728*
FRT (colour)				0.642*	0.594	0.416	0.753*	0.644*	0.507
FRT (context)					0.886**	0.092	0.722*	0.751*	0.863**
FRT (sound)						−0.126	0.846**	0.778**	0.899**
FRT (motion)							0.236	0.179	−0.057
Picture naming								0.893**	0.825**
WPM									0.816**

FRT = Feature Reality Test; WPM = word-to-picture matching; CCP = Camel and Cactus picture test. **Correlation is significant at the 0.01 level (two-tailed); *correlation is significant at the 0.05 level (two-tailed).

non-living domain (Keil, 1989). If this were the case, then semantic disruption may give rise to greater difficulty mapping between knowledge of motion and other representational features for living than non-living concepts. Knowledge of animate and inanimate motion may even depend on distinct regions of occipitotemporal cortex, as suggested by recent functional imaging experiments (Chao *et al.*, 1999).

To identify any feature-by-domain interactions in the present data, a group-level ANOVA was conducted on the patients' z-scores, using feature (four levels: colour, context, sound and motion) and animacy (two levels: animate and inanimate) as within-subject factors. There was no main effect of feature, suggesting that patients as a group found all conditions equally difficult [$F(3,33) = 1.502$, $P > 0.05$]. Animate items, however, were significantly more difficult for patients than inanimate ones [$F(3,33) = 10.344$, $P < 0.01$] and, critically, there was a significant interaction between animacy and feature [$F(3,33) = 5.011$, $P < 0.01$] due to an exaggeration of the animacy effect in the motion condition (Fig. 4).

Finally, although the inanimate items in the stimulus set are associated with significantly higher familiarity ratings (see Table 1, bottom), the domain difference in the motion condition is unlikely to be attributable to this factor alone, since a familiarity discrepancy occurs in the same direction in all four conditions, and the difference associated with items in the motion condition is neither the largest nor the most consistent.

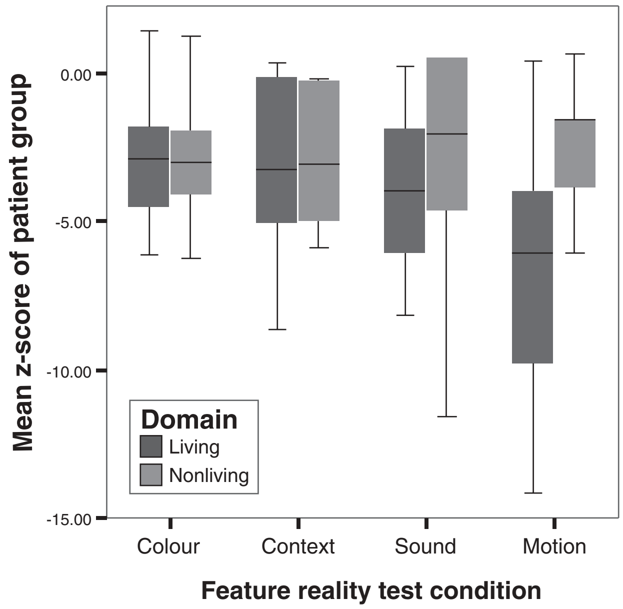

Fig. 4 Interaction between feature condition and conceptual domain produced by greater accuracy on non-living than living items in the motion condition. The representation of this effect using z-score data shows that it is not simply an exaggeration of a differential difficulty effect.

Discussion

The data reported in this paper were obtained using a novel, non-verbal battery of tests of semantic feature knowledge, designed to explore the breakdown of conceptual representations in a group of patients at a range of different stages of semantic dementia. Unlike many previous similar studies, performance on this battery requires neither spoken output nor linguistic mediation over and above an explanation of the very simple nature of the task. Consequently, the severe anomia that typically affects patients with semantic dementia, particularly those at the more severe end of the clinical spectrum, neither interfered with test administration nor confounded interpretation of their performance.

The four conditions of the FRT probed, individually, some of the fine-grained aspects of knowledge on which conceptual representations are believed to depend. In addition to the well-studied areas of object colour, location and environmental sound, the battery introduced a novel feature dimension: motion. Using these materials it was possible to probe knowledge of specific associations in a simpler and more focused manner than is possible using either the Pyramids and Palm Trees (PPT) or Camel and Cactus Test (CCT). Although both PPT and CCT can be administered in picture form, they nonetheless involve an element of online problem-solving (rendering them sensitive to deficits in executive function with or without semantic impairment), and present a relatively narrow range of semantic associations, many of which are difficult to define any more explicitly than as aspects of 'encyclopaedic knowledge'. It is also unclear, when administering either of these assessments, whether failure on any individual item reflects degraded knowledge about the index item (e.g. an elephant) or about the target and/or foil associates (e.g. a circus tent and a church).

The patterns of results obtained suggested that (i) performance on the test was closely correlated with conventional tests of semantic memory; (ii) there was significant consistency in the accuracy with which patients responded to the different feature conditions associated with an individual concept; (iii) both performance and correlation with other semantic tasks were dramatically modulated by the plausibility of the distractors; (iv) patients were able to demonstrate more knowledge about items that they were able to name than about those for which they had become anomic, even when scores were compared across item pairs matched for familiarity; and (v) knowledge about item motion did not behave in the same way as the other features or general semantic tests.

The high correlations with other semantic tests, and the finding of greater accuracy on named than unnamed items suggest, first of all, that the FRT is a valid measure of conceptual knowledge, consistent with previous studies investigating its breakdown at a fine-grained level (Bozeat *et al.*, 2000; Lambon Ralph and Howard, 2000; Miceli *et al.*, 2001; Coccia *et al.*, 2004).

The absence of performance differences between different feature conditions is a finding central to the debate about organization of fine-grained feature knowledge. Homogeneity of performance emerged not only from group-level analysis but also when scores were examined on a case-by-case basis. In the latter, only one patient (AT) showed an effect of feature on performance, though individual contrasts did not show any specific differences between feature conditions. The demonstration of consistency in both analyses is particularly striking given the two alternative forced-choice response format of the tests used, a factor that would tend to make significant item-wise associations more difficult to demonstrate. The fact that both correlation and accuracy were dramatically modulated by the plausibility of the distractor items fits well with the results of colour decision and real versus chimeric object decision tests reported by Rogers *et al.* (2003*a*, *b*): faced with these tasks, semantic dementia patients are influenced by the regularity of the item's attributes, leading to rejection of unusual-looking real items (e.g. a seahorse) and acceptance of chimeric items formed by replacing atypical features with typical ones (e.g. a giraffe with a short neck). A similar pattern has also been reported for the task of distinguishing words from non-words (Rogers *et al.*, 2004*b*), highlighting the importance of shared structure as a major determinant of stability in irregular representational domains. The present findings allow this principle to be extended to previously untested aspects of concept knowledge.

Finally, the possibility that such item-wise correspondence might be explained in terms of the generally facilitatory effect of concept familiarity on semantic task accuracy was not upheld: better performance on named than unnamed items was seen even when the comparison was restricted to familiarity-matched item pairs. This set of findings is therefore compatible with the assumption that a supramodal system underpins representation at the semantic level, as suggested by Rogers *et al.* (2004*a*).

The results also bear on the question of whether semantic dementia constitutes a unitary syndrome or includes a mixture of progressive aphasic cases, some with a lexical disorder impairing name production and single-word comprehension, and others with an additional visual associative agnosia (Mesulam, 2001). If the latter were the case, a clear mismatch would be expected between performance on language-dependent and language-independent tasks, or at the very least some degree of variability in performance on different conditions of the FRT. The homogeneity of performance of the population across task conditions reported here, together with the consistency across conditions and across tasks that depend for correct performance on intact representations of meaning, would therefore support the former interpretation.

The design of the test materials was motivated specifically by the predictions of the account proposed by Rogers *et al.* (2004*a*), the central assumption of which is the idea that semantic information is represented in terms of abstract similarity relations (i.e. independent of feature modality) among concepts. A dynamic system that receives input from multiple cortical areas, each representing a particular dimension of similarity, will come to represent the higher-order similarities between items within a complex, high-dimensional 'semantic space'. Damage to the neural structures underpinning these supramodal representations, such as the progressive anterior temporal atrophy seen in semantic dementia, would cause a disturbance of higher-order representation, leading to a characteristic pattern of errors on tests for whose performance it is critical.

The predictions of this account are of further interest in the light of the finding that category-specific deficits demonstrated on conventional measures such as naming and word-picture matching were not reflected in a differential performance on the FRT. Previous models (e.g. Warrington and Shallice, 1984) have attempted to account for such cases in terms of differential importance of sensory and non-sensory properties in the representation of living and non-living kinds. If this was the case, it might have been expected that case DW might have shown greater impairment on sensory features (colour, sound and perhaps also motion) in comparison with the purely associative feature of environmental context. This was manifestly not the case, either with DW or any of the four other patients who showed evidence of category asymmetry (IB, TW, WM and JM). This negative finding further reinforces the absence of an association between sensory knowledge and living deficits documented, using verbal materials, by Lambon Ralph *et al.* (1998).

The novel set of test materials used in this study included a subset aimed at probing knowledge of object motion, a feature that has not, to our knowledge, previously been examined in a patient population. Interestingly, motion appeared different from the other conditions in several respects. First, performance on this condition correlated with colour but not with either sound or context, while other features were significantly intercorrelated. Secondly, motion scores failed to correlate significantly with performance on standard semantic battery tests. Finally, in contrast to the sound and colour conditions, there was no apparent relationship between knowledge of an object's motion and the ability to name it.

The significance of these differences naturally includes the possibility that knowledge about motion is subject to a different set of representational constraints, in common with basic numerical knowledge, which has been shown to be selectively preserved in semantic dementia (Cappelletti *et al.*, 2001; Jefferies *et al.*, 2005). The proposal that semantics arise from organization of feature-specific representations at a supramodal level is not incompatible with this idea, as large systematic differences between domains at some peripheral level are likely to become reflected in the higher-order representational structure to which they contribute (Rogers *et al.*, 2004*a*). As previously noted, there are hints of such differences from the functional imaging and developmental literatures. Chao *et al.* (1999) found that distinct regions of occipitotemporal cortex were activated in response to living and non-living items in response to motion, suggesting a higher level of organization. Moreover, motion might intuitively be considered as more 'causally relevant' to the living than the non-living domain [an argument that has been developed by Keil (1989)]. The suggestion, however, of differences between patients in the degree to which knowledge of motion is disproportionately impaired (see Fig. 2) raises the intriguing possibility that the outlying performers (cases AN and WM) represent an anatomically atypical subgroup, with pathological involvement of posterior as well as anterior–inferior regions of temporal cortex. Patterns of impaired performance in groups of patients with other forms of temporal lobe pathology (such as Alzheimer's disease, Herpes encephalitis and stroke) together with further refinement and expansion of the battery will allow the origin of this apparent feature by domain interaction to be explored in more detail in future studies.

Acknowledgements

This work was funded by a Medical Research Council Clinician Scientist Fellowship to P.G. We are indebted to John Hodges for allowing us to study three of the patients reported here. Funding to pay the Open Access publication charges for this article was provided by the Medical Research Council.

References

Bozeat S, Lambon Ralph MA, Patterson K, Garrard P, Hodges JR. Non-verbal semantic impairment in semantic dementia. Neuropsychologia 2000; 38: 1207–15.

Bozeat S, Lambon Ralph MA, Graham K, Patterson K, Wilkin H, Rowland J, et al. A duck with four legs: investigating the structure of conceptual knowledge using picture drawing in semantic dementia. Cogn Neuropsychol 2003; 20: 27–47.

Breedin SD, Saffran EM, Coslett HB. Reversal of the concreteness effect in a patient with semantic dementia. Cogn Neuropsychol 1994; 11: 617–60.

Cappelletti M, Butterworth B, Kopelman M. Spared numerical abilities in a case of semantic dementia. Neuropsychologia 2001; 39: 1224–39.

Chao LL, Haxby JV, Martin A. Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. Nat Neurosci 1999; 2: 913–9.

Coccia M, Bartolini M, Luzzi S, Provinciali L, Ralph MAL. Semantic memory is an amodal, dynamic system: evidence from the interaction of naming and object use in semantic dementia. Cogn Neuropsychol 2004; 21: 513–27.

De Vreese LP, Faglioni P, Agnetti V. Una batteria di prove per la agnosia e afasia dei colori. Archivio di Psicologia, Neurologia e Psychiatria 1994; 55: 742–90.

Folstein MF, Robins LN, Helzer JE. The Mini-Mental State Examination. Arch Gen Psychiatry 1983; 40: 812.

Funnell E, Sheridan J. Categories of knowledge—unfamiliar aspects of living and nonliving things. Cogn Neuropsychol 1992; 9: 135–53.

Galton CJ, Patterson K, Xuereb JH, Hodges JR. Atypical and typical presentations of Alzheimer's disease: a clinical, neuropsychological, neuroimaging and pathological study of 13 cases. Brain 2000; 123: 484–98.

Garrard P, Lambon Ralph MA, Patterson K, Pratt KH, Hodges JR. Semantic feature knowledge and picture naming in dementia of Alzheimer's type: a new approach. Brain Lang 2005; 93: 79–94.

Gloor P. The temporal lobe and limbic system. New York: Oxford University Press; 1997.

Grossman M, Ash S. Primary progressive aphasia: a review. Neurocase 2004; 10: 3–18.

Hodges JR, Patterson K. Nonfluent progressive aphasia and semantic dementia: a comparative neuropsychological study. J Int Neuropsychol Soc 1996; 2: 511–24.

Hodges JR, Patterson K, Oxbury S, Funnell E. Semantic dementia. Progressive fluent aphasia with temporal lobe atrophy. Brain 1992; 115: 1783–806.

Hodges JR, Graham N, Patterson K. Charting the progression in semantic dementia: implications for the organisation of semantic memory. Memory 1995; 3: 463–95.

Hodges JR, Patterson K, Graham N, Dawson K. Naming and knowing in dementia of Alzheimer's type. Brain Lang 1996; 54: 302–25.

Hodges JR, Davies RR, Xuereb JH, Casey B, Broe M, Bak TH, et al. Clinicopathological correlates in frontotemporal dementia. Ann Neurol 2004; 56: 399–406.

Jefferies E, Bateman D, Ralph MAL. The role of the temporal lobe semantic system in number knowledge: evidence from late-stage semantic dementia. Neuropsychologia 2005; 43: 887–905.

Keil F. Concepts, kinds, and cognitive development. Cambridge, MA: MIT Press; 1989.

Kertesz A, Davidson W, McCabe P, Takagi K, Munoz D. Primary progressive aphasia: diagnosis, varieties, evolution. J Int Neuropsychol Soc 2003; 9: 710–9.

Lambon Ralph MA, Howard D. Gogi aphasia or semantic dementia. Cognitive Neuropsychology 2000; 17: 437–65.

Lambon Ralph MA, Graham KS, Ellis AW, Hodges JR. Naming in semantic dementia–what matters? Neuropsychologia 1998; 36: 775–84.

Lambon Ralph MA, Powell J, Howard D, Whitworth AB, Garrard P, Hodges JR. Semantic memory is impaired in both dementia with Lewy bodies and dementia of Alzheimer's type: a comparative neuropsychological study and literature review. J Neurol Neurosurg Psychiatry 2001; 70: 149–56.

Luzzatti C, Davidoff J. Impaired retrieval of object-colour knowledge with preserved colour naming. Neuropsychologia 1994; 32: 933–50.

McKenna P, Warrington EK. The Graded Naming Test. Windsor, Berks: NFER-Nelson; 1983.

Mesulam MM. Primary progressive aphasia. Ann Neurol 2001; 49:425–32.

Miceli G, Fouch E, Capasso R, Shelton JR, Tomaiuolo F, Caramazza A. The dissociation of colour from form and function knowledge. Nat Neurosci 2001; 4: 662–7.

Patterson K, Hodges JR. Deterioration of word meaning: implications for reading. Neuropsychologia 1992; 30: 1025–40.

Patterson K, Lambon Ralph MA, Hodges JR, McClelland JL. Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge. Neuropsychologia 2001; 39: 709–24.

Rogers T, Hodges J, Patterson K, Lambon Ralph M. Object recognition under semantic impairment: the effects of conceptual regularities on perceptual decisions. Lang Cogn Processes 2003a; 18: 625–62.

Rogers TT, Patterson K, Hodges J, Graham K. Colour knowledge in semantic dementia: it's not all black and white. J Cogn Neurosci 2003b; Supplement: 222.

Rogers TT, Lambon Ralph MA, Garrard P, Bozeat S, McClelland JL, Hodges JR, et al. Structure and deterioration of semantic memory: a neuropsychological and computational investigation. Psychol Rev 2004a; 111: 205–35.

Rogers TT, Ralph MAL, Hodges JR, Patterson K. Natural selection: the impact of semantic impairment on lexical and object decision. Cogn Neuropsychol 2004b; 21: 331–52.

Rosen HJ, Allison SC, Schauer GF, Gorno-Tempini ML, Weiner MW, Miller BL. Neuroanatomical correlates of behavioural disorders in dementia. Brain 2005; 128: 2612–25.

Snodgrass JG, Vanderwart M. Standardized set of 260 pictures—norms for name agreement, image agreement, familiarity, and visual complexity. J Exp Psychol-Human Learning and Memory 1980; 6: 174–215.

Snowden J, Goulding P, Neary D. Semantic dementia: a form of circumscribed cerebral atrophy. Behav neurol 1989; 2: 167–82.

Snowden JS, Neary D, Mann DMA, editors. Frontotemproal lobar degeneration: frontotemporal dementia, progressive aphasia, semantic dementia. London: Churchill Livingstone; 1996.

Warrington EK. The selective impairment of semantic memory. Quart J Exp Psychol 1975; 27: 635–57.

Warrington EK, Shallice T. Category specific semantic impairments. Brain 1984; 107: 829–54.

Warrington EK, James M. The visual object and space perception battery. Bury St Edmunds, UK: Thames Valley Test Company; 1991.