

# At the Edge of Semantic Space: The Breakdown of Coherent Concepts in Semantic Dementia Is Constrained by Typicality and Severity but Not Modality

Emily J. Mayberry, Karen Sage, and Matthew A. Lambon Ralph

## Abstract

■ Hub-and-spoke models of semantic representation suggest that coherent concepts are formed from the integration of multiple, modality-specific information sources with additional modality-invariant representations—most likely stored in the ventrolateral anterior temporal lobe (vATL). As well as providing the necessary computational mechanisms for the complexities of feature integration, these modality-invariant representations also license a key aspect of semantic memory—semantic-based generalization. Semantic dementia allows us to investigate this aspect of conceptual knowledge because (a) the patients have a selective and progressive semantic degradation and (b) this

is associated with profound ventrolateral ATL atrophy. Specifically, the boundaries between concepts become degraded in semantic dementia and, when tested using the appropriate materials, the patients make simultaneous under- and overgeneralization errors. We found that the rate of these errors were a function of typicality and pseudotypicality of the items as well as the severity of the patients' semantic impairment. Following the modality-invariant nature of the vATL hub representation, we also confirmed that the patients were impaired on both verbal- and picture-based versions of the same task. ■

## INTRODUCTION

Patients with semantic dementia (SD) present with a selective yet progressive impairment of semantic memory, affecting all receptive and expressive modalities (Lambon Ralph & Patterson, 2008; Bozeat, Lambon Ralph, Patterson, Garrard, & Hodges, 2000; Warrington, 1975). This selective neuropsychological pattern is associated with a particular distribution of atrophy. Volumetric analyses of atrophy and hypometabolism identify the polar and ventrolateral anterior temporal lobe (vATL) as the worst affected region (Nestor, Fryer, & Hodges, 2006; Galton et al., 2001; Mummery et al., 2000), although atrophy spreads to posterior temporal and other regions as the disease progresses (Rohrer et al., 2009). Given the distribution and increasing spread of atrophy in SD, it is important to note that convergent evidence for the importance of the ATL (especially the ventrolateral region) arises from other independent sources. For example, the same behavioral pattern is mirrored when repetitive transcranial magnetic stimulation is applied to the left or right lateral ATL. Specifically, this induces a temporary yet selective slowing of both receptive and expressive semantic tasks, for both verbal and nonverbal materials (Pobric, Jefferies, & Lambon Ralph,

2007, 2010a; Lambon Ralph, Pobric, & Jefferies, 2009). Convergent evidence for a role of the ATL in semantic cognition comes from MEG, PET, and also fMRI once various technical and design issues are taken into account (Visser, Embleton, Jefferies, Parker, & Lambon Ralph, 2010; Binder, Desai, Graves, & Conant, 2009; Marinkovic et al., 2003). The extension to fMRI studies is especially helpful, given that it offers greater spatial resolution for addressing the question of which specific parts of the ATL appear to be critical for semantic memory (the anterior-ventral and anterior-lateral areas: Binney, Embleton, Jefferies, Parker, & Lambon Ralph, 2010).

This study aims to answer one of the next generation of more specific questions about the modality-invariant hub within the “hub-and-spoke” framework (Pobric, Jefferies, & Lambon Ralph, 2010b; Patterson, Nestor, & Rogers, 2007)—what does it contribute to semantic representation? Many contemporary theories of semantic memory assume that concepts are formed from the mass, parallel action of information arising in multiple, modality-specific brain areas (Martin, 2007; Barsalou, Simmons, Barbey, & Wilson, 2003) and closely related ideas can be traced back to the classical theories of conceptualization, proposed by Wernicke and Meynert (Eggert, 1977). Such theories are intuitively appealing because there is no mystery about where semantic knowledge comes from—it derives

---

University of Manchester, UK

directly from the totality of our verbal and nonverbal experience. More importantly, modern functional neuroimaging has been able to show that multiple association regions are activated when participants are required to process the meaning of different concepts (Martin, 2007).

The next key neuroscience questions are how and where these information sources combine in order to form coherent and generalizable semantic representations. The computations required are nontrivial because semantic features combine to form concepts in complex, nonlinear ways and over an extended period of time (for a more in-depth discussion of these issues, see Lambon Ralph, Sage, Jones, & Mayberry, 2010). Different aspects of this challenge have been highlighted and discussed previously in the philosophical (Wittgenstein, 2001), cognitive science (Murphy & Medin, 1985; Smith & Medin, 1981; Rosch, 1975), and computational modeling literatures (Rogers & McClelland, 2004). Although coming from different perspectives, they have all noted that concepts are grounded in our sensorimotor and verbal experience but that coherent, generalizable concepts require some kind of additional representation or step (e.g., family resemblances, prototypes, modality-invariant representation). The computational models of Rogers et al. (2004) and Rogers and McClelland (2004) provide a formal and implemented instantiation of this general idea. In these hub-and-spoke semantic frameworks, concepts are formed from the combination of modality-specific sources of information (the spokes) with a learnt, modality-invariant representation (the hub) (for a direct test of this framework, see Pobric et al., 2010b). Rather than forming representations via a web-like set of connections between all possible modality pairings, the modality-invariant hub generates a multidimensional semantic space into which modality-specific features can be mapped. The multiple dimensions of the hub representational space are critical because they are able to code the complex, nonlinear mappings between concepts and the multitude of modality-specific features, and thus, provide a basis for conceptually based generalization (Lambon Ralph et al., 2010; Rogers et al., 2004; Rogers & McClelland, 2004).

Our working hypothesis is that distributed regions throughout the cortex continue to represent modality-specific information, whereas regions within the ventrolateral ATL provide a core neural substrate for this conceptual hub (Lambon Ralph et al., 2010; Pobric et al., 2010b; Lambon Ralph & Patterson, 2008; Patterson et al., 2007). Some of the more basic predictions from this framework are consistent with existing findings from multiple neuroscience methods. For example, (1) the ATL hub contains modality-invariant codes, and thus, this region should be involved in both verbal and nonverbal semantic tasks (Pobric et al., 2010a; Bozeat et al., 2000; Vandenberghe, Price, Wise, Josephs, & Frackowiak, 1996), and (2) modality-specific regions should still contribute to conceptual knowledge but only when the semantic task probes information relevant to that specific modality (Pobric et al., 2010b).

In a recent study, we also demonstrated evidence consistent with another prediction. A core characteristic of coherent semantic representations is that they allow us to generalize on the basis of conceptual rather than superficial similarities (e.g., we know that information about dogs should apply to Golden Retrievers and West Highland Terriers but only a subset of it to wolves, foxes, cats, and other similar creatures). Our hypothesis was that, as a result of the progressive semantic degradation, there should be a reduction in the computational power of the modality-invariant hub representation, and so SD patients should find it increasingly difficult to match items on the basis of conceptual rather than superficial similarities (see Figure 4, Lambon Ralph et al., 2010). By using a form of semantic matching-to-sample, we were able to demonstrate that this pattern does emerge in SD and that the patients increasingly make both overgeneralization (incorrectly extending category membership to noncategory exemplars) and undergeneralization (omitting category exemplars as members of the category) errors, often simultaneously (e.g., incorrectly including a fox as a dog while also failing to include a Chihuahua as a dog; Lambon Ralph et al., 2010). The present study was designed to use this novel finding and assessment method in order to investigate three key factors: (1) semantic typicality/pseudotypicality; (2) the severity of semantic impairment; and (3) the modality in which the information is presented, for instance, words versus pictures. We discuss each of these briefly, in turn.

Typicality has a long-standing history in research on semantic memory (Rosch, 1975). We explored its impact on the modality-invariant hub contribution to semantic representation for two reasons. First, our hypothesis is that these representations license the formation of concepts free from superficial similarities. As originally noted by Rosch and previously by Wittgenstein, there is considerable variation among the exemplars for any given concept. Some exemplars share many features in common and they are rated as being typical, whereas other exemplars can be superficially very different (Garrard, Lambon Ralph, Hodges, & Patterson, 2001; Rosch, 1975). Despite this, our semantic system is able to generalize our knowledge for each concept across all exemplars, that is, the system can overcome surface statistical regularities and encompass atypical exemplars within the same conceptual boundary. In the hub-and-spoke framework, the hub is critical for pulling together information for all concepts. In addition, we might expect the hub contribution to be especially important for capturing atypical exemplars within the correct boundary (and there is already some evidence for the importance of typicality in the context of SD naming performance: Woollams, Cooper-Pye, Hodges, & Patterson, 2008). Our second reason for exploring this variable is that discussions of typicality most often focus on the relationship between exemplars and a single category or concept. *We think, however, this reflects only half of the challenge for the*

*semantic system*. For accurate semantic generalization, the semantic system must also set a conceptual boundary that *excludes* noncategory exemplars (which also vary in their superficial similarity to true exemplars). There are many real-life categories that have considerable overlap (e.g., fruit vs. vegetables; tools vs. sports equipment vs. weapons) making accurate conceptual differentiation even more challenging. Our previous study found that SD patients make errors in both directions and so in this study we tested the hypothesis that undergeneralizations (missing a category exemplar) would be a function of typicality and overgeneralizations (falsely accepting a noncategory exemplar) would reflect “psuedotypicality.” Given that concept familiarity is a very important performance factor in SD and it is correlated with typicality, we ensured that the materials were matched for this variable. In addition, because a word version of the task was used, we also ensured that the materials were matched for lexical frequency.

Our second target was to explore the impact of the severity of impairment to the hub component. By investigating SD patients at different levels of semantic severity, it allows us to establish when in their decline the conceptual boundaries start to break down and how this manifests itself in terms of the balance of over- and undergeneralization errors. We also expected to see a modulation of this basic factor by typicality/psuedotypicality. Specifically, given that atypical targets and high psuedotypical foils are the most challenging for the conceptual system (because their surface characteristics are the most misleading), then these exemplars might be the first to show impairment in the milder patients. With further breakdown of the multidimensional space, however, we would expect the categorization errors to extend to more typical exemplars and less psuedotypical nonexemplars.

Our third target was modality-of-presentation. A considerable amount of past SD research has emphasized the multimodal nature of semantic impairment observed in SD (Luzzi et al., 2007; Rami, Loy, Hailstone, & Warren, 2007; Coccia, Bartolini, Luzzi, Provinciali, & Lambon Ralph, 2004; Bozeat, Lambon Ralph, Patterson, & Hodges, 2002; Bozeat et al., 2000; Lambon Ralph & Howard, 2000). If the semantic hub utilizes modality-invariant representations, then we should expect to see over- and undergeneralizations, and sensitivity to typicality/psuedotypicality for both picture and verbal materials. Our previous study only utilized picture-based materials, and thus, it is important to ensure that the patients’ decisions were not driven by a nonsemantic strategy (i.e., judging purely on the basis of visual similarity). Given the arbitrary mapping between objects’ names and their meanings, we expected to observe lower performance on the verbal than pictorial versions of the same task (Lambon Ralph & Howard, 2000) but, nevertheless, we should observe parallel influences of typicality/psuedotypicality and patient severity across both versions.

## METHODS

### Participants

Five SD patients (mean age = 67.4 years, range = 60–80 years old; mean years education = 11.3) and 10 neurologically intact control participants matched to the patients for age and years of education (mean age = 68.2 years, range = 63–75 years old, mean years education = 12.1) participated in this study. The SD patients were recruited from memory clinics in Bath and Manchester (UK) and all five presented with the characteristic symptoms of SD: selective, progressive semantic impairment with other aspects of cognitive functioning preserved. All had CT or MRI scans conducted at the time of clinical diagnosis, which showed circumscribed ATL atrophy. All five patients had more left than right ATL atrophy with the mildest patient, D. F., showing only left-sided changes and the rest with some degree of bilateral temporal lobe atrophy. The patients completed a range of background neuropsychological assessments to determine each patient’s degree of semantic impairment and preservation of other cognitive skills. Table 1 shows a summary of results from these background tests with the patients ordered in terms of semantic severity. Although some other cognitive skills were beginning to decline in the most severe case (E. T., who had very significant semantic impairment at this stage), these neuropsychological assessments reflect the standard SD pattern of selective semantic impairment.

### Materials

#### *Collection of Familiarity, Typicality, and Psuedotypicality Ratings*

Eight control participants (mean age = 68.3 years, range = 63–75) provided ratings of category members from 14 different categories (insects, vegetables, tools, fruit, cooking utensils, sports equipment, appliances, birds, household items, musical instruments, weapons, fish, furniture, and vehicles). Each control participant was given a written category name (e.g., musical instruments) followed by a list of category members (e.g., violin, guitar, didgeridoo), and was asked to make several judgments about each item: (1) *category membership*: participants were asked whether each item was or was not a member of the given category; (2) *typicality*: to try to lessen the influence of familiarity (Garrard et al., 2001), participants were asked to imagine an average of all the category members they knew and to disregard how often they encounter the items. They were prompted to take into account what they look like, where you might see them, what kinds of sounds they might make, what they are made of, what they might eat if they are living, what you might use them for, and anything else they could think of about the items. They were asked to use this average as a comparison point and indicate how close or far each item was from their imaginary category average [from 1 (*not close/not typical*) to 7 (*very close/very typical*)]; (3) *familiarity*: participants were also asked

**Table 1.** Background Measures of Semantic Impairment and Cognitive Skills

		<i>D. F.</i>	<i>M. T.</i>	<i>M. B.</i>	<i>P. L.</i>	<i>E. T.</i>
Sex		M	F	F	F	F
Age		63	60	62	72	80
Distribution of ATL atrophy		L	L > R	L > R	L > R	L > R
<i>Semantic Test Battery</i>						
64 Naming	(Max = 64)	54	45	35	22	0
64 Word–picture match	(Max = 64)	60	57	48	44	11
CCT (word)	(Max = 64)	55 <sup>a</sup>	46	40	29	14
CCT (pictures)	(Max = 64)	56 <sup>a</sup>	45	41	30	28
PPT (word)	(Max = 52)	46	44	40	n/t	32
PPT (pictures)	(Max = 52)	47	44	44	n/t	31
Synonym judgment	(Max = 96)	87	77	71	78	50
Category fluency	(Six categories)	57	67	45	26	8
Letter fluency	(F, A, S)	13	40	20	24	22
<i>General Cognitive Measures</i>						
ACE-R	(Max = 100)	78	79	72	56	43
Digit span	Forward	7	7	6	8	7
	Backward	4	5	6	5	6
RCPM	(percentile)	95%	95%	90%	95%	95%
Rey Figure Copy	(Max = 36)	36	36	33	31	23.5
<i>VOSP</i>						
Screening	(Max = 20)	20	20	19	19	18
Incomplete letters	(Max = 20)	19	20	20	19	16
Silhouettes	(Max = 20)	13	8	4	3	0
Object decision	(Max = 20)	20	20	18	15	13
Dot counting	(Max = 10)	10	10	10	10	10
Position discrimination	(Max = 20)	20	20	18	18	19
Number location	(Max = 10)	10	10	10	7	7
Cube analysis	(Max = 10)	10	10	10	9	8

CCT = Camel and Cactus Test; PPT = Pyramids and Palm Trees Test; ACE-R = Addenbrooke's Cognitive Examination (Revised); RCPM = Ravens Coloured Progressive Matrices; VOSP = Visual and Spatial Test Battery; n/t = not tested; L = left; R = right.

<sup>a</sup>Denotes the only semantic scores to fall within the normal range for these tests.

how familiar each item was to them; they were asked to disregard typicality but to think about how often they came across each item/concept (whether through reading about it, seeing it, hearing about it, thinking about it) and also how much they knew about that item. They used another 7-point Likert scale to indicate their familiarity ratings with 1 (*not familiar at all*) to 7 (*very familiar*).

In order to select foil items, we asked the participants to generate two additional sets of cross-category ratings. For potential foils, we asked them to indicate whether or not an item belonged to the target category (e.g., is a banana a member of the vegetable category?) Then they were asked to repeat the typicality rating procedure, as described above. Instead of rating the exemplar's typicality

of its own category, they were asked to pretend that the item belonged to the new category and to rate its “psuedotypicality” (e.g., consider all of the vegetables you know, and then rate how similar a banana is to your average vegetable).

### Stimuli Selection

To be considered for selection, at least 75% of the participants had to agree with our initial category placement (e.g., that both bananas and avocados are fruit) in their category judgments. For the cross-category foil judgments, items were only selected if at least 75% of the participants agreed that the exemplar did *not* belong to the novel category (e.g., that bananas are not vegetables). The final test materials comprised 33 trials. Each “trial” was made up of six choices: two category members and four nonmembers. For the category members, we selected one more typical and one atypical exemplar (e.g., an orange and an avocado for the fruit category). The four noncategory members comprised two unrelated exemplars (e.g., a bayonet and a tomahawk) and two closely related items. One of these related items was selected to be more psuedotypical of the target category than the other (e.g., sweet corn vs. mange tout). Statistical analysis of the properties for the selected items confirmed that the six sets of items (typical targets, atypical targets, psuedotypical foils, pseudo-atypical foils, and the two unrelated sets) were matched for familiarity [one-way ANOVA:  $F(5, 138) < 1$ ], that the typical targets were rated as being more typical than the atypical [typical  $M = 5.64$  vs. atypical  $M = 3.65$ :  $t(46) = 8.39$   $p < .001$ ], and that the psuedotypical foil was more “typical” of the target category than the pseudo-atypical foil [psuedotypical  $M = 1.48$  vs. pseudo-atypical  $M = 0.33$ :  $t(46) = 7.25$ ,  $p < .001$ ]. We also confirmed that the items were matched for lexical frequency using N-Watch (Davies, 2005) and the NeuroMCW website (two-tailed  $t$  test  $p > .50$  for contrasts of typical vs. atypical targets and psuedotypical vs. pseudo-atypical foils, and two-tailed  $t$  test  $p > .25$  for all contrasts of targets vs. foils.)

### Tasks

The patients and controls completed two tasks formed from these selected items. The first was a multiple-target, matching-to-sample test very similar in procedure to our previous study. By providing multiple targets and foils in each trial, it is possible for patients to make overgeneralization (selecting foils) and undergeneralization errors (missing targets) simultaneously (Lambon Ralph et al., 2010). Each trial commenced with the name of the target category appearing on the middle of a computer screen. The experimenter also read this word to the patients. On the following screen, they were shown a  $3 \times 2$  array of the six choices (see above) in either picture (color photographs) or word format. The picture and word versions were completed on different days to avoid interference.

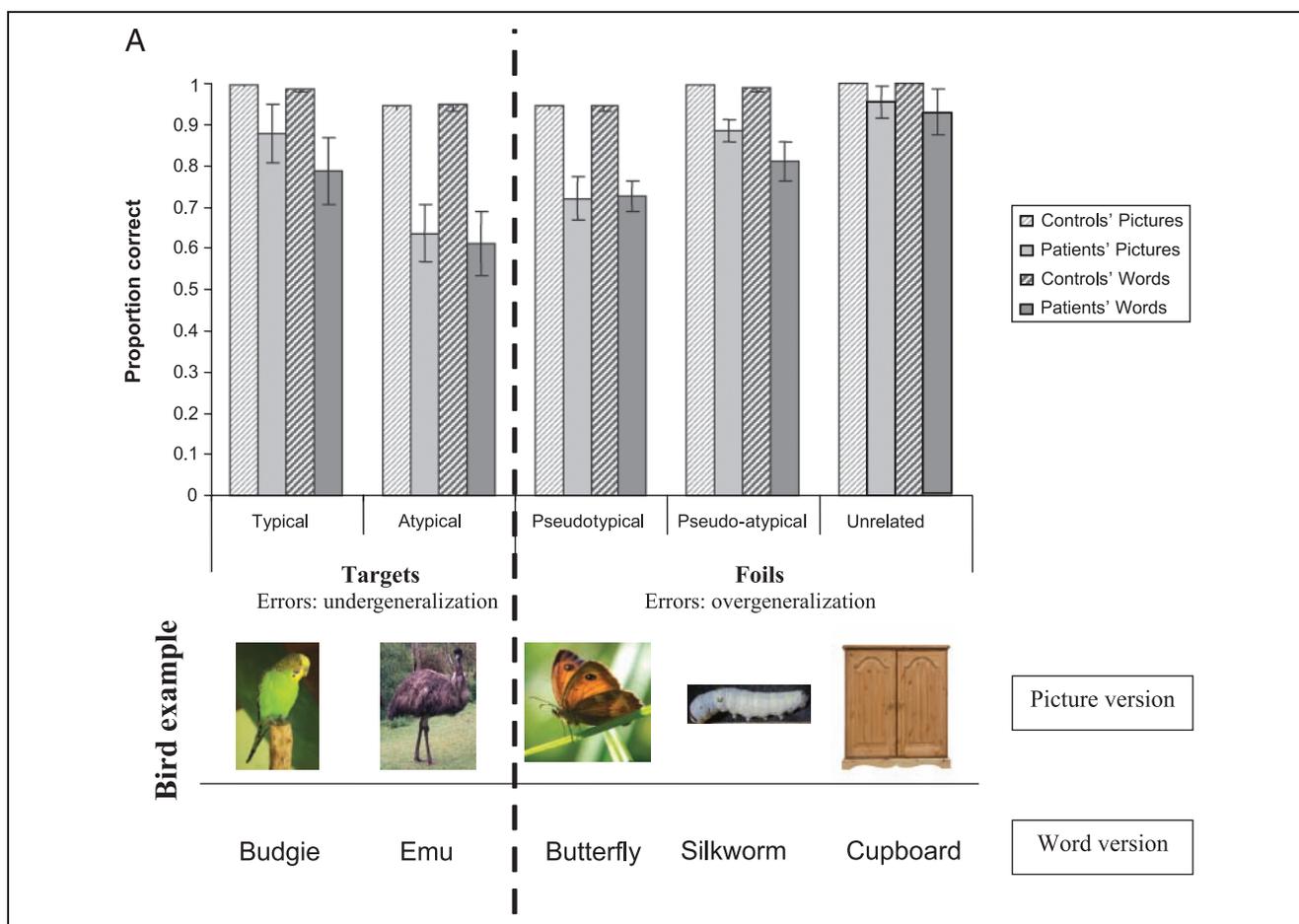
Participants were told that there were always two category exemplars in each array and that they should pick the two that they thought were most likely to be members of that category. If a patient initially only selected one item then he or she was prompted for a second choice. If a patient pointed to more than two items, we asked them to pick the two that they thought were the best category exemplars.

Although the six-choice matching-to-sample test allows us to detect and separate over- and undergeneralization errors, the patients’ categorization to each target cannot be probed directly because they are free to choose any two from the six presented items. We created, therefore, a two-alternative forced-choice (2AFC) version of the task using exactly the same materials. Six 2AFC trials were formed from the sextuplet of selected items (described above). Specifically, each target (typical or atypical) was paired with the psuedotypical foil, pseudo-atypical foil, and one of the unrelated foils (assigned randomly). This design allowed us to probe the patients’ comparison between each combination of target–foil pairs. Following our working hypothesis, we expected their accuracy to be a function of the relative typicality of the targets and foils. Specifically, we expected the typicality of the targets to have little or no effect when they were paired with unrelated items because it is possible to reject unrelated concepts with only minimal remaining semantic information. In contrast, we expected typicality to become critically important for the closely related foils: The patients’ performance should be worse when the atypical target was paired with the psuedotypical foil and better when a typical target was paired with a pseudo-atypical foil. The procedure for the 2AFC task was as follows. The category name was presented at the top of the computer screen and the two choices were shown at the bottom. Again, verbal and nonverbal (picture) versions of the task were administered on separate occasions. All written names (sample and choices) were read aloud by the experimenter. The participants were instructed to pick only one of the two alternatives.

## RESULTS

### Six-Alternative Forced-choice Task (6AFC)

Figure 1A displays the mean control and patient accuracy on the 6AFC matching-to-sample task. This is broken down into their accuracy for correctly accepting the different target types (typical and atypical) and correctly rejecting the different foil types (psuedotypical, pseudo-atypical, and unrelated) for the picture and word versions. The results show that patients performed as expected: (i) On both the picture and word versions of the task, the patients performed significantly worse than the controls [pictures:  $F(4, 44) = 45.0$ ,  $p < .001$ ; words:  $F(4, 44) = 16.5$ ,  $p < .001$ ]; (ii) there was also an interaction between choice type and participant group [pictures:  $F(4, 44) =$



**Figure 1.** (A) Mean control and semantic dementia performance on the picture and word 6AFC matching-to-sample test. The data are subdivided into the picture and the word versions of the task. In each trial (e.g., “which of these is a bird?”), the targets included one typical (e.g., budgie) and one atypical (e.g., emu) exemplar. There were also three types of foil: “pseudotypical” (e.g., butterfly), “pseudo-atypical” (e.g., silkworm), and two unrelated (e.g., cupboard, rocking chair). The SD patients made a combination of two types of simultaneous, selection error: (i) *undergeneralization*—failures to pick target exemplars (especially of the atypical exemplars) and also (ii) *overgeneralization*—incorrect selection of nontarget items (particularly pseudotypical foils). Error bars indicate standard error. (B) Individual semantic dementia performance on picture and word versions of the 6AFC matching-to-sample test. Individual patient data from the five semantic dementia patients are shown in order from least to most severe (left to right). Error bars indicate standard error.

19.9,  $p < .001$ ; words:  $F(4, 44) = 8.04, p < .001$ ; analyzed in more detail below]; and (iii) there was no interaction between materials and choice type [ $F(4, 16) = 1.29, p = .32$ ], indicating a very similar pattern of results for the word and picture materials as would be expected from the breakdown to modality-invariant representations.

More detailed analyses were conducted on the patients' results. There was a trend for performance on the word version to be worse than for the picture version [ $F(1, 4) = 5.81, p = .07$ ], although this effect was carried primarily by the most severe patient, E. T. [when her data were excluded then there was no modality effect:  $F(1, 3) = 3.06, p = .18$ ], and there was no interaction between modality and choice type [ $F(4, 12) < 1$ ]. As expected, there was a main effect of choice type [ $F(4, 16) = 15.3, p < .001$ ]. We tested and confirmed our specific predictions with regard to graded effects of typicality/pseudotypicality (see Introduction: targets: atypical > typical; foils: pseudotypical > pseudo-atypical > unrelated) with planned

paired-samples  $t$  tests. For both the picture and the word versions, patients were better at correctly accepting typical than atypical targets [picture:  $t(4) = 9.56, p = .001$ ; words:  $t(4) = 3.27, p = .03$ ]. In the picture version, patients were worse at rejecting the pseudotypical than pseudo-atypical foils [ $t(4) = 3.86, p = .02$ ]. In the word version, the patients performed equally well on the pseudotypical and pseudo-atypical foils [ $t(4) = 1.26, p = .28$ ]. In both the picture and word versions, the pseudotypical foils [pictures:  $t(4) = 5.20, p = .007$ ; words:  $t(4) = 3.12, p = .04$ ] and pseudo-atypical foils [picture:  $t(4) = 3.47, p = .03$ ; words:  $t(4) = 3.00, p = .04$ ] were more difficult to reject than the unrelated items. On both the word and picture versions of the task, all patients simultaneously undergeneralized (missed a target) and overgeneralized (incorrectly selected a foil) on some of the trials. Specifically, out of 34 picture trials, the patients simultaneously under- and overgeneralized in 11.2 trials, on average (D. F.: 6 trials; M. T.: 11 trials; M. B.: 10 trials; P. L.: 16 trials;

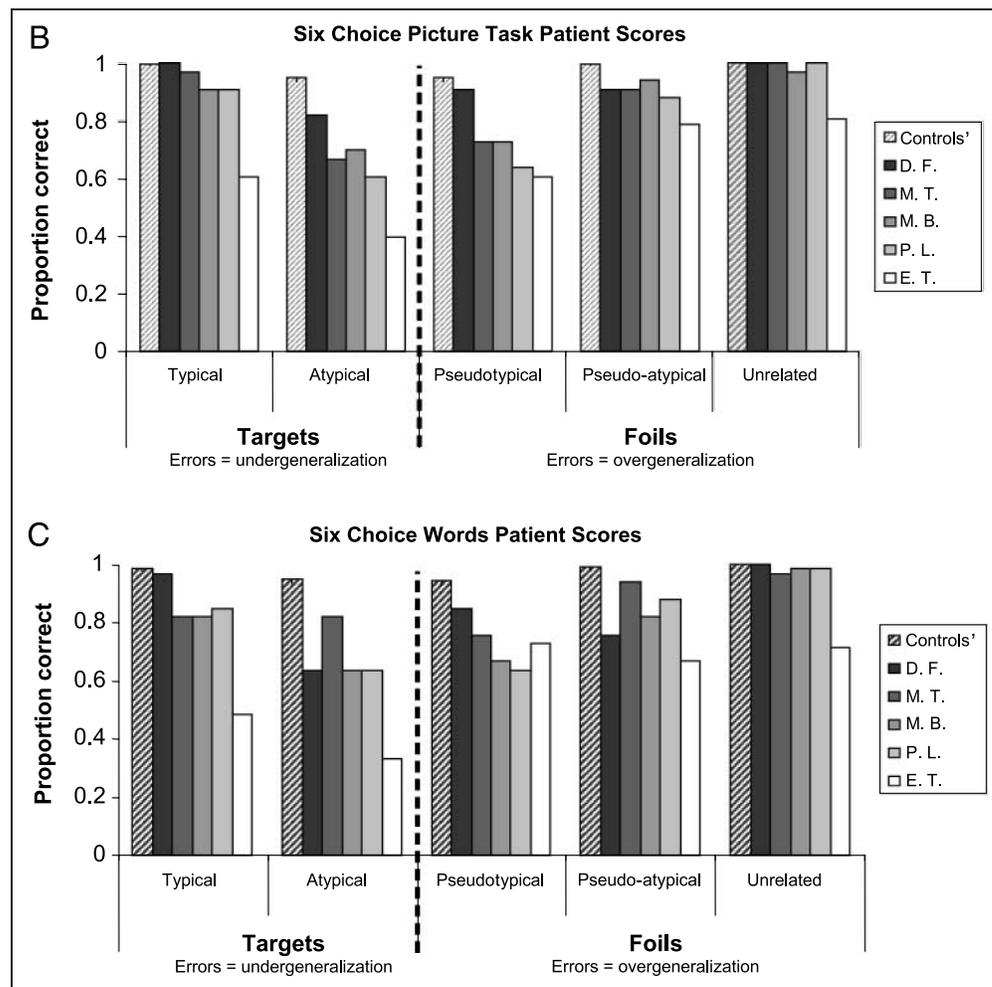
and E. T.: 13 trials). Out of the 34 word trials, the patients simultaneously under- and overgeneralized in 14.6 trials, on average (D. F.: 11 trials; M. T.: 13 trials; M. B.: 16 trials; P. L.: 13 trials; and E. T.: 20 trials). This is striking given that, within a trial, the patients are considering only one concept at a time and yet make both kinds of errors.

### Individual Performance and Influence of Semantic Severity

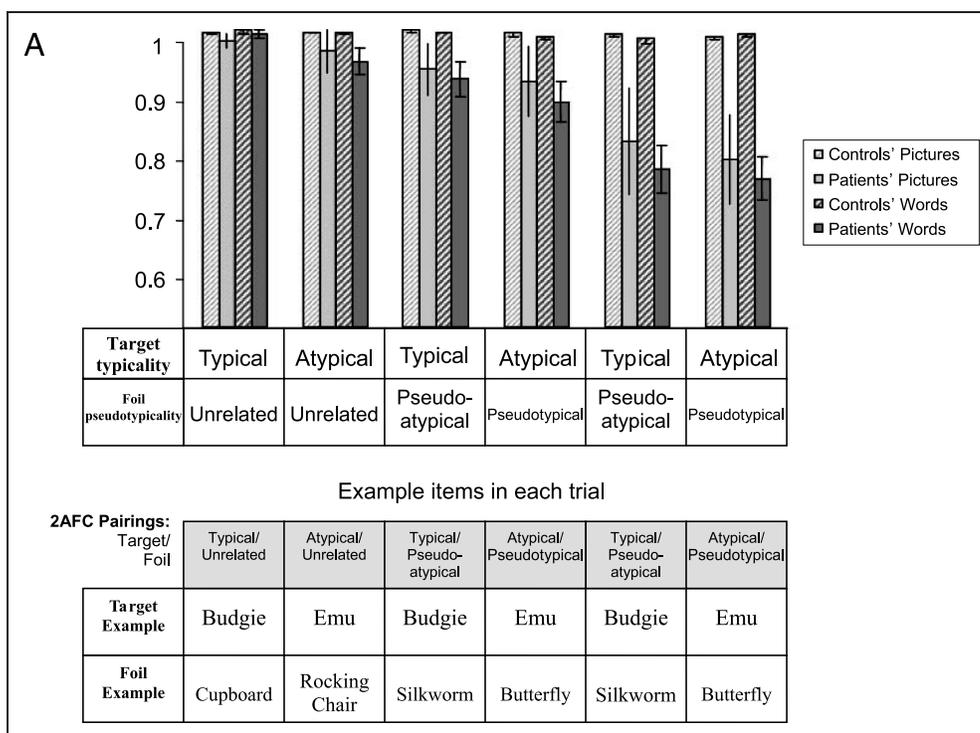
Figure 1B (picture task) and Figure 1C (word task) show the individual patient scores for each type of target and foil. The patients are ordered by semantic severity with the least severe patient, D. F., farthest to the left (next to the control average), and E. T., the most severe case, rightmost. For the typical targets, D. F. (the mildest patient) performed within the control range in both the picture and the word versions of the task, whereas the others performed outside of the normal range and their performance declined in approximate severity order. All patients except the most severely impaired (E. T.) were able to rule out the unrelated items with high accuracy. Scores on the atypical targets and the two more closely related foils types were outside the normal range for all patients.

In order to investigate the effect of semantic severity on patients' ability to accept typical and atypical targets, and to reject unrelated, pseudo-atypical and pseudo-typical foils, we fitted a linear regression to the patient scores for each type of item using a composite semantic score (which encompasses 6 semantic tasks: naming, word-to-picture matching, semantic associations in word and picture format, synonym judgment, and category fluency; this single composite semantic score—the first principal component of the variance among the 6 tests—accounts for 92% of the variance). Looking at the individual patient scores in Figure 1B, it is clear that for unrelated foils, the slope between D. F. and P. L. is nearly flat despite their wide range of severities. E. T. pulls this down but the regression slope is still very close to zero for the unrelated foils ( $\beta = 0.073$ ) and also for the pseudo-atypical foils ( $\beta = 0.049$ ). The typical targets have a steeper slope ( $\beta = 0.149, p = .016$ ), and as predicted, so did the atypical targets ( $\beta = 0.150, p = .009$ ) and the pseudo-typical foils ( $\beta = 0.103, p = .057$ ). Looking at Figure 1C, we can see that the same general pattern of results exists for the word version of the task although the data are noisier (probably due to the arbitrary mappings between words and their meanings). Because the patient case-series runs from mild

Figure 1. (continued)



**Figure 2.** (A) Mean accuracy for controls and patients on the 2AFC matching-to-sample test. Control and patient averages for the different trial types on the 2AFC version of the test. Trial types are ordered from most difficult category judgments to least difficult category judgments from left to right. In each trial, a “target,” which was either *typical* or *atypical* of its category, was paired with a “foil,” which was either (1) from an *unrelated* category, or (2) from a close category and rated as *pseudotypical* in relation to the target category or (3) from a close category and rated as being *pseudo-atypical* in relation to the target category. An example of the word version (using the target category birds) is shown below. Typ = typical; Atyp = atypical. Error bars indicate standard error. (B) Individual patient



data for the 2AFC picture and word version of the task. Individual patient scores contrasted with control averages for the different trial types on the 2AFC version of the test. Trial types are in order from most difficult category judgments to least difficult category judgments from left to right. In each trial, a “target,” which was either typical or atypical of its category, was paired with a “foil,” which was either (1) from an *unrelated* category, or (2) from a close category and rated as *pseudotypical* in relation to the target category or (3) from a close category and rated as being *pseudo-atypical* in relation to the target category. Typ = typical; Atyp = atypical. Error bars indicate standard error.

(D. F.) to very severe (E. T.), we can also see that the group pattern holds across all patients but the interaction between severity of impairment and typicality/pseudotypicality is attenuated in the most severe case, E. T., where a more generalized impairment across all conditions begins to emerge.

### Two-alternative Forced-choice Task

Figure 2A shows the control and the patient mean scores on the picture and word versions of the 2AFC test. Again, the patients performed exactly as expected. There was a three-way interaction between target typicality, foil pseudotypicality, and participant group [ $F(2, 22) = 11.4, p < .001$ ]. There was also a main effect of target typicality [typical > atypical:  $F(1, 3) = 75.2, p = .003$ ], of foil pseudotypicality [unrelated > pseudo-atypical > pseudotypical:  $F(2, 6) = 16.8, p = .003$ ], and an interaction between these two factors [ $F(2, 6) = 31.3, p = .001$ ]. Figure 2A shows that, as expected, control performance stayed at ceiling across all trial types but there was a gradation of difficulty for the patients as the foils became more pseudotypical and the targets became less typical. The patient data were then analyzed in more detail. As with the 6AFC task, performance was slightly better (though not significantly better and still impaired) on the picture than word versions of the task [ $F(1, 3) = 8.18, p = .065$ ]. Planned paired-samples  $t$  tests were conducted and confirmed our prediction with

regard to the relative importance of target typicality (typical vs. atypical) at each level of foil pseudotypicality. When the foils were unrelated, there was no difference between performance on the high versus low typicality target trials [pictures:  $t(4) < 1$ ; words:  $t(3) = 2.32, p = .10$ ]. With pseudo-atypical foils, the difference between typical and atypical targets approached significance in the picture version [ $t(4) = 2.39, p = .075$ ] and was significant in the word version [ $t(3) = 6.12, p = .009$ ]. When paired with pseudotypical foils, there was a larger and significant difference between typical and atypical targets [pictures:  $t(4) = 7.20, p = .002$ ; words:  $t(3) = 10.1, p = .002$ ].

### Individual Scores and the Influence of Semantic Severity

Figure 2B shows the individual patient scores for the different trial types. The trial types are ordered by increasing difficulty from left to right (as in the previous figure), and the patients are again ordered in terms of semantic severity. E. T., the most semantically impaired patient, was excluded from the analyses of the word version because she performed at chance level during the first 50 trials (E. T. scored 20/50).

All patients performed very well on trials in which a typical target was paired with an unrelated foil (e.g., asked to determine whether a budgie or a cupboard was a bird) in both the word and picture versions of the test. However,

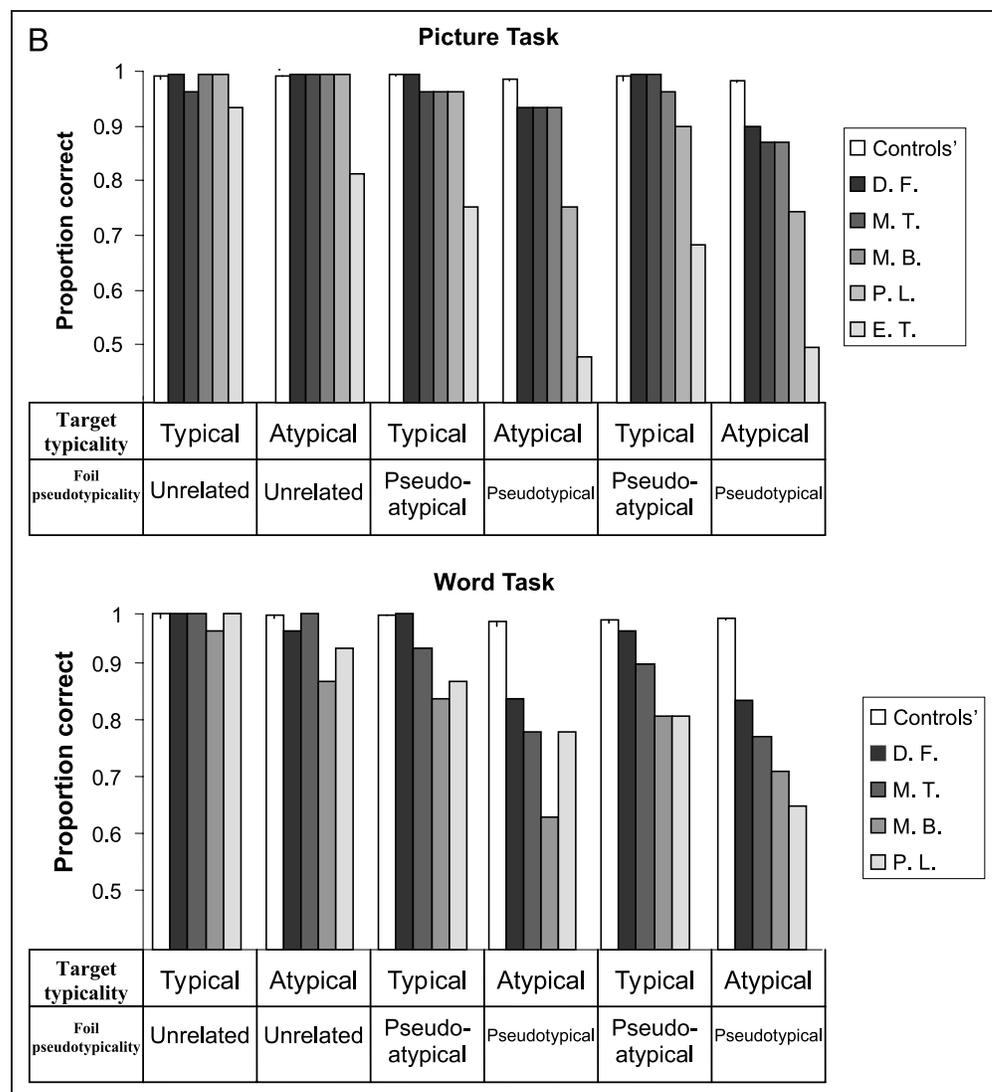
there appeared to be a gradation of difficulty as the targets became less typical and the foils became closer to the target category boundaries (more psuedotypical), and as the patients became more severe. To look at the impact of semantic severity in more detail, we computed linear regressions of the generalized semantic severity measure (see above) against each trial type (the various combinations of target typicality and foil psuedotypicality). We found that semantic severity became a much better predictor as the targets became less typical and the foils became more psuedotypical (summarized in Table 2). Parallel effects were observed in the picture and word versions but were more pronounced in the word task.

## DISCUSSION

This study used a novel form of semantic matching-to-sample to explore the specific contribution that the ATL modality-invariant hub makes to the rest of the hub-and-spoke semantic system (Pobric et al., 2010b; Patterson et al., 2007). This is essential in providing the necessary

computational mechanisms for the many complexities involved in the integration of multimodal sources of information into a coherent concept (see Introduction and Lambon Ralph et al., 2010). An ATL contribution to semantic memory has been established through neuropsychology, neuroimaging, computational modeling, and repetitive transcranial magnetic stimulation [rTMS] (e.g., Binney et al., 2010; Lambon Ralph et al., 2010; Pobric et al., 2007, 2010b; Visser et al., 2010; Lambon Ralph et al., 2009; Patterson et al., 2007; Rogers et al., 2004), and at least two subregions within the ATL seem to be critical in this role (the anterior basal temporal area and the anterior lateral region; Binney et al., 2010). However, there is currently only limited information on the exact function and role of the modality-invariant hub within semantic representation—which was the focus of this study. Specifically, we explored how this representational hub contributes to the computation of coherent concepts and investigated three key parameters: typicality/psuedotypicality, modality-of-presentation, and severity of semantic impairment.

Figure 2. (continued)



**Table 2.** Linear Regressions Relating Semantic Severity to Performance on the 2AFC Semantic Task

<i>Trial Type</i>		<i>Pictures</i>						<i>Words</i>					
<i>Target Type</i>	<i>Foil Type</i>	$R^2$	$\beta$	<i>SE</i>	<i>Std. <math>\beta</math></i>	<i>t</i>	<i>p</i>	$R^2$	$\beta$	<i>SE</i>	<i>Std. <math>\beta</math></i>	<i>t</i>	<i>p</i>
Typical	Unrelated	.43	0.018	0.012	0.66	1.52	.23	.07	0.006	0.017	0.26	0.38	.74
Atypical	Unrelated	.71	0.069	0.025	0.83	2.71	.073	.30	0.046	0.049	0.55	0.93	.45
Typical	Pseudo-atypical	.79	0.088	0.026	0.89	3.35	.044	.75	0.094	0.038	0.87	2.47	.13
Atypical	Pseudo-atypical	.89	0.187	0.039	0.94	4.83	.017	.22	0.091	0.038	0.87	2.47	.13
Typical	Pseudotypical	.91	0.125	0.023	0.95	5.35	.013	.90	0.113	0.027	0.95	4.13	.054
Atypical	Pseudotypical	.91	0.161	0.029	0.96	5.621	.011	.99	0.130	0.004	0.99	30.79	.001

Our first hypothesis was that as the hub representation breaks down, the most vulnerable items will be those immediately on either side of the “conceptual boundary.” This is because the hub system is crucial for bringing together information from a range of modalities and representing the core semantic structure of a concept independently from any particular surface features or characteristics (Lambon Ralph et al., 2010; Rogers et al., 2004). As predicted, we found that the degrading conceptual boundary could be defined, at least in part, by the item’s typicality, for both target exemplars as well as for noncategory items (which fall on the other side of the conceptual boundary). Specifically, we found that the SD patients were more accurate at accepting typical than atypical exemplars as targets and found it harder to reject pseudotypical noncategory exemplars than semantically related or unrelated foils.

The second factor we investigated in the current study was how a patient’s level of semantic impairment affected their degree of conceptual breakdown. As expected, we found that early in the progression of the disorder, patients began to make a small (but significant) number of errors on the atypical targets and the pseudotypical foils, which rely the most on the modality-invariant representations, but that they performed at normal or near-normal levels on the more typical targets and less pseudotypical foils. This pattern of typicality-driven performance held across our patient case-series, albeit at varying overall levels of accuracy. It was only in the most severe case, E. T., that this pattern was less clear, with a smaller difference between her performance on the more typical versus atypical targets and between the more pseudotypical versus pseudo-atypical foils. There are at least two possible explanations for E. T.’s less differentiated performance on the task. First of all, it is possible that E. T.’s semantic system is degraded to a point where she is unable to access very much deep conceptual or even surface knowledge about a concept (although her performance on all but the two-choice word version of the task was above chance, so she clearly had some remaining knowledge of the concepts). On the other hand, it is possible that E. T.’s performance reflects a progression of atrophy into other brain

regions beyond the ATL. As noted in the Introduction, several studies have demonstrated that as SD progresses, atrophy extends to brain regions beyond the most anterior portion of the temporal lobes (Brambati et al., 2009; Rohrer et al., 2008, 2009).

Because the representational hub is proposed to be modality-invariant in nature (e.g., Lambon Ralph et al., 2010; Pobric et al., 2010b; Patterson et al., 2007; Rogers et al., 2004), we explored the SD patients’ performance on this novel semantic matching-to-sample test in both verbal and picture-based formats. We found that, as expected, all patients were impaired on both verbal and non-verbal versions of the task, and the effects of typicality on target exemplars and foils were the same in both modalities. Overall, the patients were slightly less accurate on the word than picture version. Following previous computational and neuropsychological explorations of this issue (Rogers et al., 2004; Lambon Ralph & Howard, 2000), we interpret this relative difference in terms of the mapping type: The relationship between verbal labels and concepts is entirely arbitrary and this presents a much greater computational challenge than the quasi-systematic mapping between visual forms and meaning. As a consequence, damage to amodal semantic units leads to relatively worse performance on verbal than picture tasks.

As well as adding further evidence to the notion that the ATL is implicated in semantic memory (Simmons & Martin, 2009; Lambon Ralph & Patterson, 2008; Patterson et al., 2007), this study begins to provide some answers to the question of what specific role this region plays in semantic cognition. Other studies have begun to tackle some of the other key questions about the ATL and its contribution to semantic cognition. These include: (1) which specific areas of the ATL are implicated in the semantic hub; and (2) what are the relative contributions of modality-specific sources of information (the spokes) and the ATL hub to conceptualization? A recent fMRI study (Binney et al., 2010) begins to answer the first of these questions by providing convergent evidence with studies of SD patients (Jefferies, Patterson, Jones, & Lambon Ralph, 2009; Nestor et al., 2006; Galton et al., 2001) and rTMS studies (Lambon Ralph et al., 2009; Pobric et al., 2007). Binney et al. (2010)

showed that the regions of the ATL maximally affected in SD are also the same as those activated when healthy participants perform a semantic task: Specifically, both lines of research indicate a crucial role of the anterior inferior temporal gyrus, anterior fusiform gyrus, and the anterior superior temporal sulcus. A recent rTMS study (Pobric et al., 2010b) addressed the second question above and confirmed that the ATL hub is involved in representing all concepts (they found that living and non-living, manipulable and nonmanipulable items were equally impaired following rTMS to the ATL). In addition, this study showed that the “spokes” (association areas coding modality-specific sources of information) are involved in representing concepts but only if the spoke contains information relevant for that concept. Specifically, Pobric et al. (2010b) applied rTMS to the inferior parietal lobes (known to be important for praxis; Buxbaum & Saffran, 2002) and, in contrast to the pan-modality effect after ATL stimulation, found that nonliving items were more affected than living items and, perhaps most critically, highly manipulable items were more affected than less manipulable items.

In conclusion, several recent studies have advanced our understanding of the ATL contribution to semantic processing. Specifically, these investigations have provided information about which ATL regions are involved in semantic representation (e.g., which particular parts of the anterior temporal lobes function as a modality-invariant hub; Binney et al., 2010), what their roles might be (Pobric et al., 2010b), and what the hub itself might contribute (e.g., coherent, generalizable concepts; Lambon Ralph et al., 2010). The current study took this one step further by exploring three key factors which affect modality-invariant representations within the hub. First of all, as the hub deteriorates, the typicality of the targets and the pseudotypicality of foils become crucially important for correct categorization. Atypical exemplars (which fall near the boundary of their category) and pseudotypical noncategory items (which are on the perimeter of the category boundary) were most susceptible to under- and overgeneralization. Strikingly, the patients made these two types of error simultaneously to the same concept, suggesting an active degradation of the core semantic structure. This pattern held for all levels of severity except the most severe patient. Finally, the pattern of results was similar across the picture and word versions of the task, providing further evidence that the hub is modality-invariant.

### Acknowledgments

We thank the patients, their families, and the control participants who took part in this study. We also thank Prof. Alistair Burns, Prof. Roy Jones, Dr. Andrew Larner, and Dr. Roland Zahn for referring patients to us. E. J. M. was supported by an Overseas Research Student (ORS) Award and the University of Manchester. This project was supported by an MRC programme grant to M. A. L. R. (G0501632).

Reprint requests should be sent to Prof. Matthew A. Lambon Ralph, Neuroscience and Aphasia Research Unit (NARU), Zochonis Building, School of Psychological Sciences, University of Manchester, Oxford Road, Manchester, M13 9PL, UK, or via e-mail: matt.lambonralph@manchester.ac.uk.

### REFERENCES

- Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, 7, 84–91.
- Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, 19, 2767–2796.
- Binney, R. J., Embleton, K., Jefferies, E., Parker, G. J. M., & Lambon Ralph, M. A. (2010). The inferolateral aspects of the anterior temporal lobe are crucial in semantic memory: Evidence from a novel direct comparison of distortion-corrected fMRI, rTMS and semantic dementia. *Cerebral Cortex*, 20, 2728–2738.
- Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia*, 38, 1207–1215.
- Bozeat, S., Lambon Ralph, M. A., Patterson, K., & Hodges, J. R. (2002). When objects lose their meaning: What happens to their use? *Cognitive, Affective & Behavioural Neuroscience*, 2, 236–251.
- Brambati, S. M., Rankin, K. P., Narvid, J., Seeley, W. W., Dean, D., Rosen, H. J., et al. (2009). Atrophy progression in semantic dementia with asymmetric temporal involvement: A tensor-based morphometry study. *Neurobiology of Aging*, 30, 103–111.
- Buxbaum, L. J., & Saffran, E. M. (2002). Knowledge of object manipulation and object function: Dissociations in apraxic and nonapraxic subjects. *Brain and Language*, 82, 179–199.
- Coccia, M., Bartolini, M., Luzzi, S., Provinciali, L., & Lambon Ralph, M. A. (2004). Semantic memory is an amodal, dynamic system: Evidence from the interaction of naming and object use in semantic dementia. *Cognitive Neuropsychology*, 21, 513–527.
- Davies, C. J. (2005). N-Watch: A program for deriving neighbourhood size and other psycholinguistic statistics. *Behavior Research Methods*, 37, 65–70.
- Eggert, G. H. (1977). *Wernicke's works on aphasia: A source-book and review*. The Hague: Mouton.
- Galton, C. J., Patterson, K., Graham, K., Lambon-Ralph, M. A., Williams, G., Antoun, N., et al. (2001). Differing patterns of temporal atrophy in Alzheimer's disease and semantic dementia. *Neurology*, 57, 216–225.
- Garrard, P., Lambon Ralph, M. A., Hodges, J. R., & Patterson, K. (2001). Prototypicality, distinctiveness and intercorrelation: Analyses of the semantic attributes of living and nonliving concepts. *Cognitive Neuropsychology*, 18, 125–174.
- Jefferies, E., Patterson, K., Jones, R. W., & Lambon Ralph, M. A. (2009). Comprehension of concrete and abstract words in semantic dementia. *Neuropsychology*, 23, 492–499.
- Lambon Ralph, M. A., & Howard, D. (2000). Gogi aphasia or semantic dementia? Simulating and assessing poor verbal comprehension in a case of progressive fluent aphasia. *Cognitive Neuropsychology*, 17, 437–466.
- Lambon Ralph, M. A., & Patterson, K. (2008). Generalization and differentiation in semantic memory: Insights from semantic dementia. In A. Kingstone & M. Miller (Eds.), *The year in cognitive neuroscience* (pp. 61–76). New York: Annals of the NY Academy of Sciences.

- Lambon Ralph, M. A., Pobric, G., & Jefferies, E. (2009). Conceptual knowledge is underpinned by the temporal pole bilaterally: Convergent evidence from rTMS. *Cerebral Cortex*, *19*, 832–838.
- Lambon Ralph, M. A., Sage, K., Jones, R. W., & Mayberry, E. J. (2010). Coherent concepts are computed in the anterior temporal lobes. *Proceedings of the National Academy of Sciences, U.S.A.*, *107*, 2717–2722.
- Luzzi, S., Snowden, J. S., Neary, D., Coccia, M., Provinciali, L., & Lambon Ralph, M. A. (2007). Distinct patterns of olfactory impairment in Alzheimer's disease, semantic dementia, frontotemporal dementia, and corticobasal degeneration. *Neuropsychologia*, *45*, 1823–1831.
- Marinkovic, K., Dhond, R. P., Dale, A. M., Glessner, M., Carr, V., & Halgren, E. (2003). Spatiotemporal dynamics of modality-specific and supramodal word processing. *Neuron*, *38*, 487–497.
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, *58*, 25–45.
- Mummery, C. J., Patterson, K., Price, C. J., Ashburner, J., Frackowiak, R. S. J., & Hodges, J. R. (2000). A voxel based morphometry study of semantic dementia: The relation of temporal lobe atrophy to cognitive deficit. *Annals of Neurology*, *47*, 36–45.
- Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316.
- Nestor, P. J., Fryer, T. D., & Hodges, J. R. (2006). Declarative memory impairments in Alzheimer's disease and semantic dementia. *Neuroimage*, *30*, 1010–1020.
- Patterson, K., Nestor, P. J., & Rogers, T. T. (2007). Where do you know what you know? The representation of semantic knowledge in the human brain. *Nature Reviews Neuroscience*, *8*, 976–987.
- Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2007). Anterior temporal lobes mediate semantic representation: Mimicking semantic dementia by using rTMS in normal participants. *Proceedings of the National Academy of Sciences, U.S.A.*, *104*, 20137–20141.
- Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2010a). Amodal semantic representations depend on both anterior temporal lobes: Evidence from repetitive transcranial magnetic stimulation. *Neuropsychologia*, *48*, 1336–1342.
- Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2010b). Category-specific versus category-general semantic impairment induced by transcranial magnetic stimulation. *Current Biology*, *20*, 964–968.
- Rami, L., Loy, C., Hailstone, J., & Warren, J. (2007). Odour identification in frontotemporal lobar degeneration. *Journal of Neurology*, *254*, 431–435.
- Rogers, T. T., Lambon 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*, 205–235.
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.
- Rohrer, J. D., McNaught, E., Foster, J., Clegg, S. L., Barnes, J., Omar, R., et al. (2008). Tracking progression in frontotemporal lobar degeneration: Serial MRI in semantic dementia. *Neurology*, *71*, 1445–1451.
- Rohrer, J. D., Warren, J. D., Modat, M., Ridgway, G. R., Douiri, A., Rossor, M. N., et al. (2009). Patterns of cortical thinning in the language variants of frontotemporal lobar degeneration. *Neurology*, *72*, 1562–1569.
- Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General*, *104*, 192–233.
- Simmons, W. K., & Martin, A. (2009). The anterior temporal lobes and the functional architecture of semantic memory. *Journal of the International Neuropsychological Society*, *15*, 645–649.
- Smith, E. E., & Medin, D. L. (1981). *Categories and concepts/Edward E. Smith and Douglas L. Medin*. Cambridge, MA: Harvard University Press.
- Vandenberghe, R., Price, C., Wise, R., Josephs, O., & Frackowiak, R. S. J. (1996). Functional-anatomy of a common semantic system for words and pictures. *Nature*, *383*, 254–256.
- Visser, M., Embleton, K. V., Jefferies, E., Parker, G. J. M., & Lambon Ralph, M. A. (2010). The anterior temporal lobes and semantic memory clarified: Novel evidence from distortion-corrected spin-echo EPI fMRI. *Neuropsychologia*, *48*, 1689–1696.
- Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, *27*, 635–657.
- Wittgenstein, L. (2001). *Philosophical investigations: The German text, with a revised English translation 50th anniversary commemorative edition*. Oxford: Wiley-Blackwell.
- Woollams, A. M., Cooper-Pye, E., Hodges, J. R., & Patterson, K. (2008). Anomia: A doubly typical signature of semantic dementia. *Neuropsychologia*, *46*, 2503–2514.