



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Semantic diversity accounts for the "missing" word frequency effect in stroke aphasia

Citation for published version:

Hoffman, P, Rogers, TT & Lambon Ralph, MA 2011, 'Semantic diversity accounts for the "missing" word frequency effect in stroke aphasia: insights using a novel method to quantify contextual variability in meaning', *The Journal of Cognitive Neuroscience (JoCN)*, vol. 23, no. 9, pp. 2432-2446.
<https://doi.org/10.1162/jocn.2011.21614>

Digital Object Identifier (DOI):

[10.1162/jocn.2011.21614](https://doi.org/10.1162/jocn.2011.21614)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

The Journal of Cognitive Neuroscience (JoCN)

Publisher Rights Statement:

© Hoffman, P., Rogers, T. T., & Lambon Ralph, M. A. (2011). Semantic diversity accounts for the "missing" word frequency effect in stroke aphasia: insights using a novel method to quantify contextual variability in meaning. *Journal of Cognitive Neuroscience*, 23(9), 2432-2446. 10.1162/jocn.2011.21614

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



Semantic diversity accounts for the “missing” word frequency effect
in stroke aphasia: Insights using a novel method to quantify
contextual variability in meaning

Paul HOFFMAN^{1*}

Timothy T. ROGERS²

Matthew A. LAMBON RALPH¹

¹Neuroscience and Aphasia Research Unit, School of Psychological Sciences, University of
Manchester, UK

²University of Wisconsin-Madison

* Correspondence to:

Dr. Paul Hoffman, School of Psychological Sciences, Zochonis Building, University of
Manchester, Oxford Road, Manchester, M13 9PL, UK

Tel: +44 (0) 161 275 7336

Fax: +44 (0) 161 275 2873

Email: paul.hoffman@manchester.ac.uk

Acknowledgements

We are indebted to the patients and their carers for their generous assistance with this study.
The research was supported by an MRC programme grant (G0501632) and an NIHR senior
investigator grant to MALR.

Abstract

Word frequency is a powerful predictor of language processing efficiency in healthy individuals and in computational models. Puzzlingly, frequency effects are often absent in stroke aphasia, challenging the assumption that word frequency influences the behaviour of any computational system. To address this conundrum, we investigated divergent effects of frequency in two comprehension-impaired patient groups. Patients with semantic dementia (SD) have degraded conceptual knowledge as a consequence of anterior temporal lobe atrophy and show strong frequency effects. Patients with multimodal semantic impairments following stroke (semantic aphasia; SA), in contrast, show little or no frequency effect. Their deficits arise from impaired control processes that bias activation towards task-relevant aspects of knowledge. We hypothesised that high frequency words exert greater demands on cognitive control because they are more semantically diverse – they tend to appear in a broader range of linguistic contexts and have more variable meanings. We developed a new measure of semantic diversity based on latent semantic analysis that reflected the variability of a word's meaning across different contexts. Frequency, but not diversity, was a significant predictor of comprehension in SD, whereas diversity was the best predictor of performance in SA. Most importantly, SA patients *did* show typical frequency effects but only when the influence of diversity was taken into account. These results are consistent with the view that higher frequency words place higher demands on control processes, so that when control processes are damaged the intrinsic processing advantages associated with higher-frequency words are masked.

Introduction

A very common finding in studies of healthy language processing is that high frequency (HF) words are processed more quickly and accurately than low frequency (LF) words. This is true, for example, in picture naming (Oldfield, 1966), reading aloud (Forster & Chambers, 1973) and lexical decision (Balota & Chumbley, 1984), and is attributed to HF words having more robust representations in the language system (Plaut, McClelland, Seidenberg, & Patterson, 1996), higher resting levels of activation (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Dell, 1989) or higher prior probabilities (Norris, 1994). On this basis, one might expect neurological damage to be most detrimental to LF words and indeed this is the case in some aphasic groups such as patients with semantic dementia (SD), who show large frequency effects in receptive and expressive language tasks (Bozeat, Lambon Ralph, Patterson, Garrard, & Hodges, 2000; Funnell, 1995; Jefferies, Patterson, Jones, & Lambon Ralph, 2009; Lambon Ralph, Graham, Ellis, & Hodges, 1998). In stroke aphasia, however, frequency effects are often absent or, in occasional extreme cases, even reversed such that LF words are processed more accurately than HF words (Crutch & Warrington, 2005; Hoffman, Jefferies, & Lambon Ralph, submitted; Jefferies & Lambon Ralph, 2006; Marshall, Pring, Chiat, & Robson, 2001; Nickels & Howard, 1995; Warrington & Cipolotti, 1996; Warrington & Shallice, 1979). These findings present a challenge to most current models and theories of language processing which have typically focused on explaining the ubiquitous processing advantages shown by HF words (e.g., Forster & Chambers, 1973; Norris, 1994) and so offer little leverage on understanding the absence or reversal of such effects in some forms of aphasia.

This study addresses the conundrum by exploiting a striking and consistent divergence of frequency effects observed in two comprehension-impaired patient groups (Jefferies & Lambon Ralph, 2006): semantic dementia and stroke-aphasic patients with multimodal semantic impairments (referred to here as semantic aphasia or SA). The hypothesis we wish to test begins with a consideration of the different cognitive and neural mechanisms that support semantic cognition and how these can break down in neurological disorders. As many others have suggested, semantic abilities are supported by two interacting systems: a system of knowledge representation that allows

us to decode the properties of objects and the meanings of words as we encounter them; and a system of control that guides the retrieval of task-relevant semantic information in a given situation or context (Badre & Wagner, 2002; Jefferies & Lambon Ralph, 2006; Rogers & McClelland, 2004; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997). The representational component of semantic cognition has been studied for many years and, whilst it is still subject to lively debate, there is a growing consensus that various temporal lobe sites and other widely-distributed regions in posterior cortex support semantic representation (Hickok & Poeppel, 2007; Martin, 2007; Patterson, Nestor, & Rogers, 2007; Pobric, Jefferies, & Lambon Ralph, 2010; Simmons & Martin, 2009; Visser, Jefferies, & Lambon Ralph, 2010). More recently, it has become clear from functional neuroimaging and neuropsychological studies that the control processes that bring relevant information to the fore and prevent irrelevant aspects of knowledge from intruding are situated in ventrolateral prefrontal and inferior parietal regions (Hoffman, Jefferies, & Lambon Ralph, in press; Nagel, Schumacher, Goebel, & D'Esposito, 2008; Noonan, Jefferies, Corbett, & Lambon Ralph, 2010; Rodd, Davis, & Johnsrude, 2005; Thompson-Schill et al., 1997; Wagner, Pare-Blagoev, Clark, & Poldrack, 2001), which are associated with cognitive control more generally (Garavan, Ross, Li, & Stein, 2000; Hon, Epstein, Owen, & Duncan, 2006; Peers et al., 2005).

Cognitive control is critical for healthy semantic cognition because each object or word is associated with a vast quantity of information and only a small subset of this knowledge is helpful at any one time. If, for example, one hears the word “bark” during a conversation with a friend, bringing to mind the image of an angry dog will aid comprehension if your companion is telling you about their new pet but not if you are discussing the tree at the bottom of their garden (Badre & Wagner, 2002; Corbett, Jefferies, Ehsan, & Lambon Ralph, 2009; Jefferies & Lambon Ralph, 2006; Koenig & Grossman, 2007; Noonan et al., 2010). The same is also true of non-ambiguous words that vary in meaning across different contexts—consider, for instance, the various meanings of the word “on” in the phrases “on the wall” (i.e., attached to a vertical surface), “on the table” (supported by a horizontal surface), and “on TV” (appearing in the view screen of the television). In cases where there is a great deal of information or multiple different potential meanings associated with a given word or object,

only some of which is useful in the current context, control processes are thought to play a pivotal role in ensuring that only the right information comes to mind at the right time (Thompson-Schill et al., 1997).

Semantic representation and semantic control processes can be independently impaired in neurological patients. Semantic dementia (SD) is a progressive neurodegenerative disorder in which atrophy focussed on the anterior temporal lobes accompanies a pan-modal deterioration of semantic knowledge (Bozeat et al., 2000; Snowden, Goulding, & Neary, 1989). SD patients show high levels of consistency when their impoverished knowledge store is probed in different ways, such that different test formats and stimulus modalities yield similar results (Bozeat et al., 2000; Garrard & Carroll, 2006). Knowledge about properties that are shared by many semantically-related items is less vulnerable than item-specific knowledge, which is evident in the over-generalised responses these patients make in picture naming (e.g., horse → “animal”; bottle → “container” ; Woollams, Cooper-Pye, Hodges, & Patterson, 2008). This pattern of multimodal comprehension impairment, high consistency across tasks and relative preservation of superordinate information is consistent with damage to a central store of amodal semantic representations (Bozeat et al., 2000; Rogers et al., 2004). SD patients do not have difficulties using their remaining knowledge appropriately, suggesting that control processes are largely intact. Indeed, clinical observation suggests that SD patients can use their remaining conceptual knowledge in highly flexible ways to cope in everyday life.

Some stroke patients present with multimodal comprehension impairments that are superficially rather similar to those seen in SD—for instance, patients with transcortical sensory aphasia. Closer examination, however, has revealed that these impairments are qualitatively different, reflecting impaired control processes rather than a damaged semantic store (Jefferies & Lambon Ralph, 2006). Such patients show little consistency across different tasks and are highly sensitive to task demands. They perform well in semantic tasks where the correct response is strongly driven by the stimulus but not when the response is more open-ended (Corbett, Jefferies, & Lambon Ralph, 2009; Jefferies, Patterson, & Lambon Ralph, 2008; Noonan et al., 2010). For example, their picture naming improves markedly when they are provided with phonological cues that direct them toward

the correct response and their comprehension is improved by sentence cues that reduce ambiguity in word meaning. In the absence of such cues, they often make associative errors which indicate that some item-specific information has been activated but they have been unable to settle on the correct response (e.g., squirrel → “nuts”; lorry → “diesel”). This pattern, which we term semantic aphasia (SA), reflects an inability to bias processing correctly within the semantic system, so that irrelevant information often becomes highly active and has an undue influence on both verbal and non-verbal behaviour. SA patients have lesions to either left inferior frontal cortex or to the angular gyrus (Noonan et al., 2010), consistent with neuroimaging studies of semantic control. Such studies consistently reveal activation in these areas for controlled semantic processing (Badre & Wagner, 2002; Noonan, Jefferies, Visser, & Lambon Ralph, submitted; Rodd et al., 2005; Thompson-Schill et al., 1997; Zempleni, Renken, Hoeks, Hoogduin, & Stowe, 2007).

The distinction between semantic representations that code knowledge and semantic control processes that govern their proper activation can account for many of the observed differences between SD and SA (e.g., Jefferies & Lambon Ralph, 2006). The differential effect of word frequency in these two groups remains puzzling, however. The typical finding in healthy individuals, i.e., more efficient processing of HF words, is neatly captured by the basic, intuitive principle that “practice makes perfect”. The more frequently a word is encountered, the better the system learns to respond to it and the more quickly it can be processed. This intuition is well captured by the learning algorithms operating in connectionist models of mental processes (e.g., Seidenberg & McClelland, 1989). Stimuli that are presented more often in training are learned more quickly by such models and exert a greater influence on the connection weights, so that processing of such items is more resistant to damage. Accordingly, frequency effects are seen in connectionist models of reading (Plaut et al., 1996), lexical decision (Plaut, 1997) and semantic representation (Rogers & McClelland, 2004). In other computational formalisms, frequency effects are built in directly by giving HF words higher resting levels of activation (Coltheart et al., 2001; Dell, 1989; Stemberger, 1985) or higher prior probabilities (e.g., Norris, 1994).

The powerful influence of frequency on comprehension and production in SD is a natural consequence of the basic computational principle that increased exposure leads to more robust representations. Because HF words develop stronger semantic representations during learning, they are less susceptible to the general degradation of the semantic store (Rogers et al., 2004). Lambon Ralph et al. (1998) proposed two additional ways in which the frequency effect might be exaggerated in SD. First, it is possible that continued exposure to words during progression of the disease might help to counteract the breakdown of knowledge and this would benefit HF words to a greater extent (Welbourne & Lambon Ralph, 2005). Second, it may be that HF words have more detailed semantic representations, because we have more opportunities to experience them and associate them with other information. Put simply, we may know more about them and this richness may protect them from degradation of the knowledge store to some extent.

Why then are frequency effects not observed in SA patients? If SA arises from a deficit in semantic control processes, the relative executive demands of comprehending different types of word are likely to be the critical factor in determining which words are processed correctly. One might argue that HF words require less executive regulation because of their intrinsic representational strength, but this view again makes it difficult to understand why frequency effects are not observed in SA. In this article we propose a different hypothesis: that HF words typically place *higher* demands on the semantic control system than LF words and that this difference is sufficient to counteract the expected processing advantage for HF words in patients with SA. We propose that when a word is processed by the semantic system, it automatically activates information about its associations, characteristics and potential meanings, derived from prior experience. To understand the significance of the word in the current situation, semantic control systems bias processing toward the subset of this knowledge that is relevant and away from other information that is currently unhelpful (Badre & Wagner, 2002; Corbett, Jefferies, Ehsan et al., 2009; Jefferies & Lambon Ralph, 2006; Rodd et al., 2005; Thompson-Schill et al., 1997). HF words may pose particular challenges for semantic control because they have a large number of lexical associates (Nelson & McEvoy, 2000), appear in a large number of linguistic contexts (Adelman, Brown, & Quesada, 2006) and have a large number of

meanings or senses (Reder, Anderson, & Bjork, 1974). A LF word like *spinach*, for example, almost always occurs in the context of eating and cooking where it invariably refers to a particular variety of green vegetable. In contrast, a HF word like *dog* can be used in a wide variety of ways—beyond referring to a certain familiar type of four-legged animal, it can denote perseverance (“The detective will dog your footsteps”), poor physical attractiveness (“What a dog!”), a state of decrepitude (“He’s really gone to the dogs”), a state of pathetic loyalty (“Tony Blair is America’s dog”), and so on. Here the intended meaning of the word varies strongly with context, so achieving the right meaning for a given usage may require a degree of cognitive control. Or, consider a more abstract HF noun like *chance*. This word can be used in a broad range of contexts: for example, to denote a situation governed by luck (“It’s down to chance”), an opportunity that may arise in the future (“I’ll do it when I get a chance”) or a risky option (“Take a chance”). Though related, the precise meaning of *chance* is different in each case and must be inferred from the context. Thus whenever we encounter the word *chance* we might expect it to activate a rather complex set of semantic information, some of which will be irrelevant to the specific context in which it currently appears. Our hypothesis is that SA patients are unable to regulate this activation and select the appropriate information needed to comprehend the word.

Throughout this article, we use the term *semantic diversity* to denote the degree to which a given word is associated with a variety of different senses, meanings, or linguistic contexts. In the next section we describe a novel method for quantifying semantic diversity using latent semantic analysis, a computational technique for extracting representations of word meaning from large text corpora (Landauer & Dumais, 1997). We then use this measure to test the hypothesis that HF words tend to be more semantically diverse than LF words. In the following section we investigate whether a word’s semantic diversity predicts the likelihood of comprehending it in SA and SD. We demonstrate that semantic diversity is strongly associated with comprehension in SA but not in SD, consistent with the view that semantic control processes are damaged in SA but not SD. Critically, we also show that the missing word frequency effect in SA emerges once semantic diversity is taken into account.

Part 1: Measuring semantic diversity

In this section we describe our approach to quantifying semantic diversity and investigate its relationship to other lexical-semantic variables. These relationships are important for evaluating our prediction that HF words are more semantically diverse than LF words. In designing a measure of semantic diversity, we considered two closely-linked sources of variability in semantic information. The first was variability in the core meaning of a word. Some words have multiple unrelated meanings (e.g., *bark*) and many more have a number of different senses (e.g., *twist one's ankle* vs. *twist the truth*). This type of variation can be assessed by consulting dictionary definitions (e.g., Rodd, Gaskell, & Marslen-Wilson, 2002). We also, however, wanted to take into account variation in the linguistic contexts in which a word could appear. Irrespective of changes in their meanings, some words occur in a wider range of discourse topics than others and we consider this to be an important contributor to diversity in their semantic representations.

Previous work has attempted to estimate this kind of variation by asking participants to explicitly rate different words on a Likert scale (Schwanenflugel & Shoben, 1983). Such measures are, however, relatively coarse and it is not clear whether explicit ratings accurately reflect the true extent of contextual variability of different kinds of words. A more direct approach is to analyse word occurrence data in large text corpora (Adelman et al., 2006). For instance, Adelman et al. measured what they termed “contextual diversity” by counting up the number of documents containing a given word within a large text corpus. This approach is admirably objective but has two limitations. First, this measure, though not identical to word frequency, is intimately confounded with it: more frequent words must, by definition, occur in more documents. To tease apart influences of word frequency and semantic diversity, we sought a measure that can capture semantic diversity independent of word frequency. Second, Adelman et al.’s contextual diversity measure misses a key component of semantic diversity: it does not take into account similarity of meaning among the documents in which a word is found. As such the measure does not differentiate words that appear in documents with largely similar meanings from those occurring in documents with quite dissimilar meanings. Thus, we

sought a measure that would take into account similarity of meaning among the documents containing a given word.

To achieve these goals, we developed an approach based on latent semantic analysis (LSA), one of a number of techniques that use patterns of word co-occurrence to construct high-dimensional semantic spaces (Griffiths, Steyvers, & Tenenbaum, 2007; Landauer & Dumais, 1997; Lund & Burgess, 1996). The basic premise underpinning these techniques is that the linguistic context in which a word appears provides useful information about its meaning, so that words appearing in similar contexts are likely to have similar meanings (Harris, 1954; Wittgenstein, 1953). For instance, one clue to the fact that *dog* and *cat* have somewhat similar meanings is that we tend to say similar things about them: Both dogs and cats can be fed, petted, let out, collared, spayed, cuddled, and so on—so these words tend to co-occur with similar sets of other words. Firth (1957) summarised this approach as “you shall know a word by the company it keeps”.

Modern computing power has allowed this theory to be applied to large corpora of real-world language with considerable success. Landauer and Dumais’s (1997) LSA method divides a large text corpus into a number of contexts, where each context is a sample of text on a particular topic. LSA then tabulates a co-occurrence matrix indicating which words appear in which contexts. In this matrix, each word is represented as a vector, with each element of the vector corresponding to its frequency of occurrence in a particular context. Underlying structure in the co-occurrence matrix is extracted using a data-reduction technique called singular value decomposition (mathematically similar to principal components analysis), which extracts *latent* higher-order relationships between words, based on their patterns of co-occurrence. The result of these computations is a set of vectors in which each vector corresponds to a particular word in the corpus. Each vector can be thought of describing the location of a word in a high-dimensional semantic space, in which the topography of the space is determined by the underlying statistical structure of word co-occurrence. Thus, words that appear in similar linguistic environments have similar vectors and are close to one another in the space.

Most LSA studies make use of these word vectors to predict aspects of human semantic knowledge. For example, one can compare the similarity structure of the vectors to human judgements

of semantic relatedness. So, if *cat* and *dog* have similar vectors then they should also be judged by subjects to be similar in meaning. Measures of semantic relatedness derived in this way accurately predict human performance in a number of semantic tasks (Bullinaria & Levy, 2007; Landauer & Dumais, 1997; Recchia & Jones, 2009). Other studies have used the vector similarities between individual words to investigate the size of the semantic neighbourhoods of particular words (Buchanan, Westbury, & Burgess, 2001; Durda & Buchanan, 2008; Mirman & Magnuson, 2008; Shaoul & Westbury, 2006). In the present study, our goal was somewhat different. Rather than determine the semantic similarity of single words, we were interested in the relationships between the various linguistic contexts that contain a particular word. To investigate this, we made use of a second set of vectors generated by the LSA process but rarely used to investigate semantic representation. In this set, each vector corresponds to a linguistic context and not to a single word. The vectors describe the locations of contexts in the high-dimensional semantic space. So, just as two words with similar vectors are assumed to have similar meanings, two contexts with similar vectors should have similar semantic content.

To determine the semantic diversity of a particular word, we computed the pairwise similarity of all of the contexts containing the word and computed their average similarity to one another. When the contexts were very similar to one another, this suggested that the word was strongly associated with a particular type of context, and hence associated with a restricted set of semantic information and likely to have a highly constrained meaning. When the contexts associated with a given word were quite dissimilar to one another, this suggested that the word was used in a broader range of situations and was more likely to be associated with a diverse set of semantic knowledge and have a more variable, context-dependent meaning. In addition, the average similarity among associated contexts is in principle completely independent of the word's frequency. This measure therefore meets our two desiderata: it is, at least in principle, independent of word frequency and it captures the degree to which the different contexts containing a given word are similar in meaning.

Implementation of LSA

Our implementation of LSA used the written text portion of the British National Corpus (British National Corpus Consortium, 2007). This contains over 87 million words, ensuring that even LF words appeared in a substantial number of contexts. It is made up of 3125 separate documents but as many of these were very long they were sub-divided into separate contexts of 1000 words in length.¹ There were 87,375 contexts in total. LSA was carried out using the General Text Parser (Giles, Wo, & Berry, 2002). Words were only included if they appeared in at least 40 contexts and occurred at least 50 times in the corpus as a whole (38,456 words met these criteria). Prior to singular value decomposition, values in the matrix were log-transformed and each word's values divided by its entropy in the corpus (given by the formula $-\sum(p \log p)$ over all of the contexts in which it appears, where p equals context frequency divided by overall corpus frequency). These standard transformations were performed to reduce the influence of very high frequency function words whose patterns of occurrence are not relevant in generating the semantic space (Landauer & Dumais, 1997). Singular value decomposition was then used to produce a solution with 300 dimensions, which is in the region of optimum dimensionality for LSA models.

Testing the semantic representations of individual words

The semantic representations of LSA spaces are typically assessed by comparing the similarity of the vectors of word pairs with human judgements of semantic relatedness (Bullinaria & Levy, 2007; Landauer & Dumais, 1997; Recchia & Jones, 2009). Though our ultimate goal in this study was to assess relationships between contexts rather than between individual words, it was important to confirm that our implementation of LSA accurately captured the semantic relationships between words. We assessed this following the approach of Landauer and Dumais (1997), who tested their model with the synonym judgement section of the Test of English as a Foreign Language. Its accuracy (64%) was comparable to that of non-native English speakers applying to American universities. In

¹ We also generated LSA spaces and derived semantic diversity values after dividing the corpus into 250-word and 100-word blocks. The results of these analyses were highly correlated with the 1000-word model but as the 1000-word model was the best predictor of comprehension in SA patients and word recognition in healthy participants, we only report results from this model.

our case, we used the 96-item synonym judgement test, described in more detail in Part 2. Each trial consisted of a probe word and three choices, one of which had a similar meaning to the probe. We calculated the cosine between the vector representing the probe word and each of the three choices (the standard method for measuring distance between words in the LSA space), and judged the model to have “selected” whichever choice had the highest cosine with the probe. It chose the correct option on 82% of trials (significantly higher than the expected chance level of 33.3%: Binomial $p < 0.00001$). Native English speakers score 95% on average on this test. However, many of the model’s errors occurred on trials with the highest error rates and longest RTs in human participants so the model’s judgements of word similarity were a good approximation of human performance.

Calculation of semantic diversity

To determine the semantic diversity of a word, we analysed the similarity structure of the context vectors rather than the vectors corresponding to individual words. For each word, for which a semantic diversity value was required, we took all contexts in which it appeared and calculated the cosine between each pair of context vectors in the set. We then took the mean of these cosines to represent the average similarity between any two contexts containing the word.² The resulting values were log-transformed and their signs were reversed so that words appearing in more diverse contexts had higher values. Examples of semantic diversity values generated in this way are given in Table 1.³ The highest values belonged to words that could occur in almost any linguistic context. The lowest values were obtained for words with highly specific meanings that were likely to occur in a restricted range of contexts. *Gastric*, for example, only appeared in medical discussions relating to the digestive system and these contexts were represented as very similar to one another in the LSA space.

-Table 1 around here-

² In fact, the process of calculating cosines for *all* pairwise combinations of contexts would have been computationally prohibitive for very high frequency words that appeared in many thousands of contexts. To make the process more tractable, we analysed a maximum of 2000 contexts for each word. When a word appeared in more than 2000 contexts we randomly selected 2000 for the analysis.

³ We intend to make the full set of values available in a subsequent publication.

Relationship between semantic diversity and other psycholinguistic variables

A key requirement of our working hypothesis about semantic impairment in SA is that HF words tend to be more semantically diverse than LF words. Other researchers have further suggested that semantic diversity is also related to imageability, as abstract words appear in a wider range of contexts than do concrete words (Saffran, Bogyo, Schwartz, & Marin, 1980; Schwanenflugel & Shoben, 1983). We investigated the relationship between these variables and semantic diversity in a set of 4933 words for which imageability ratings were available from the MRC Psycholinguistic Database (Coltheart, 1981) and log-frequency values were available from the BNC. We also obtained the log of the number of senses given for each word in Wordnet 2.1 (<http://wordnet.princeton.edu>), a lexical database that lists word meanings (homonyms and related senses). We included this variable to gauge the extent to which the semantic diversity measure, based on the contextual usage of words, was related to polysemy as explicitly assessed by psycholinguists.

Results: The full correlation matrix is given in Table 2 and the relationship of each variable with semantic diversity depicted in Figure 1. There were strong correlations between each of the psycholinguistic variables and semantic diversity: HF words, highly abstract words and words with numerous meanings were all likely to appear in a more diverse set of contexts. These three variables were entered into a simultaneous linear regression model to predict semantic diversity. Together they accounted for 43% of the variance in semantic diversity and all contributed independently to the regression model (see Table 3).

Figure 1 also shows that, although semantic diversity is correlated with the number of senses listed in WordNet 2.0, it also captures variation in word meaning not expressed by this measure. Although words judged to have many senses invariably have high semantic diversity, words explicitly judged by psycholinguists to have few senses varied strongly in their semantic diversity. Figure 2 shows a contour plot illustrating the mean semantic diversity at all points of the frequency and imageability spectrum. The plot indicates an interaction between the two factors: Imageability covaries strongly with semantic diversity in the middle of the frequency range, but less strongly for the highest and lowest frequency words.

-Tables 2 and 3 and Figures 1 and 2 around here-

These findings show that semantic diversity could play an important role in accounting for the lack of frequency effect in SA patients: as suggested by our working hypothesis, HF words tend to be more semantically diverse and so may require a greater degree of semantic control to arrive at the correct meaning. This relationship between semantic diversity and log word frequency was, however, far from perfect (see Figure 1). Only the very highest frequency words invariably showed high levels of diversity, presumably because we use them so often that they crop up in all linguistic contexts. Nor was it the case that higher frequency words were more diverse simply by virtue of appearing in more contexts. In general, words appearing in the same number of contexts could have quite different diversity values, depending on the similarity of the contexts to one another (e.g., *turmoil* and *monarch* both appeared in 482 contexts but *monarch*'s diversity value of 1.29 was much smaller than *turmoil*'s 1.89). Across the full set of words, the number of contexts each word appeared in only accounted for 36% of the variance in semantic diversity values. This suggests that word frequency and semantic diversity can potentially account for independent variance in performance of lexico-semantic tasks.

Finally, it is worth noting that our results support another long-standing contention in the literature, that less imageable / more abstract words are also more semantically diverse. This in turn raises the possibility that previously documented differences in the processing of abstract versus concrete words might actually be attributable to differences in semantic diversity. In the next section, we investigate how word frequency, imageability and semantic diversity contribute to performance of one semantic task—synonym judgment—in SA and SD patients.

Part 2: Using semantic diversity to account for comprehension impairments

In this section, we tested one prediction of our working hypothesis: that verbal comprehension impairments in SA, but not SD, are influenced by semantic diversity, with more diverse words eliciting poorer performance because they place greater demands on semantic control. We investigated verbal comprehension using a synonym judgement task in which words were presented in with minimal disambiguating context. We assumed that this form of presentation would place

maximum demands on semantic control mechanisms, because presentation of the stimulus word would elicit activation of a wide set of semantic information relating to all of the possible contexts in which the word could be encountered. In the absence of any strong context, semantic control processes would play a key role in directing activation towards the currently relevant aspects of knowledge. This regulation process is inherently more difficult for semantically diverse words that are associated with a broader set of knowledge. We were particularly interested in the peculiar tendency of SA patients to be unaffected by word frequency. We have shown that frequency is positively correlated with diversity, raising the possibility that the frequency effect in this patient group is negated by the higher semantic diversity of HF words.

Participants

Thirteen aphasic stroke patients were recruited from stroke clubs and speech and language therapy services in Manchester, UK. Every case had chronic impairment from a CVA at least a year previously. They were classified as SA patients on the basis of criteria previously described by Jefferies and Lambon Ralph (2006). All showed cross-modal comprehension impairments in both picture and word versions of a semantic association test (Camel and Cactus Test) (Bozeat et al., 2000). We compared these patients with thirteen SD patients recruited from memory clinics in Bath, Cambridge, Liverpool and Manchester, UK, who were matched closely to the SA patients for performance on the synonym judgement test. All fulfilled published diagnostic criteria for SD (Neary et al., 1998) and scanning in each case indicated circumscribed bilateral atrophy, centred on the anterior temporal lobes. Synonym judgement data for some of these patients have been reported previously by Jefferies et al. (2007; 2009).

Demographic information and scores on background neuropsychological tests are given in Table 4. Semantic processing was assessed using the Cambridge semantic battery (Bozeat et al., 2000), which tests knowledge of 64 living and non-living items across verbal and non-verbal tasks. Both groups were impaired to a similar extent on all tests. Three subtests from the VOSP battery (Warrington & James, 1991) were used to examine visuospatial functioning, which was largely intact in both groups. The only significant differences between groups emerged on forward and backward

digit span (Wechsler, 1987) and Raven's Coloured Progressive Matrices (Raven, 1962). The lower scores of the SA group on these tests reflect impaired attentional and executive functions in this group, in line with their poor executive control over semantic information (Corbett, Jefferies, Ehsan et al., 2009; Jefferies & Lambon Ralph, 2006; Noonan et al., 2010). In contrast, the SD group scored within the normal range on these tests.

-Table 4 around here-

Synonym judgement test

Verbal comprehension was assessed using a 96-trial synonym matching test (Jefferies et al., 2009). Each trial featured a probe word and three choices, one of which had a similar meaning to the probe. The test manipulated the frequency and imageability of the probes, with two frequency bands (mean frequency = 128 vs. 4.6 counts per million in the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993)) and three bands of imageability (mean imageability = 275 vs. 452 vs. 622; values from MRC Psycholinguistic Database (Coltheart, 1981)). Frequency and imageability were varied orthogonally, such that there were 16 trials in each of the six frequency-by-imageability conditions. In addition to manipulating the probes in this fashion, the choice words on each trial belonged to the same frequency and imageability bands as the probe. During testing, the probe and three choices were presented in a written format and were also read aloud by the examiner. Patients were asked to select the choice with a similar meaning to the probe and were encouraged to guess if unsure of the correct response.

Results

For every patient, mean accuracy was computed separately for each level of frequency (2 levels) and imageability (3 levels). Mean performance for each patient group is shown in Figure 3. A 2 x 3 x 2 ANOVA was used to analyse these results, with frequency and imageability included as within-subjects factors and patient group as a between-subjects variable. This revealed main effects of frequency ($F(1,24) = 21.6, p < 0.001$) and imageability ($F(2,48) = 60.4, p < 0.001$). There were also

significant interactions between frequency and group ($F(1,24) = 19.9, p < 0.001$) and imageability and group ($F(1,24) = 3.55, p < 0.05$). While SD patients showed a robust frequency effect ($t(12) = 5.28, p < 0.001$), there was no overall frequency effect in the SA group ($t(12) = 0.18$). However, both groups did show a significant imageability effect when examined separately (SD: $F(2,24) = 50.4, p < 0.001$; SA: $F(2,24) = 24.2, p < 0.001$). The interaction arose because the SA group showed significant differences between all three levels of imageability ($t(12) > 2.38, p < 0.05$) while in the SD group there was no difference between the high and medium imageability conditions ($t(12) = 1.26, p > 0.2$), though both of these conditions produced higher accuracy levels than for the low imageability words ($t(12) > 7.45, p < 0.001$). The main ANOVA also revealed a significant frequency by imageability interaction ($F(2,48) = 5.98, p < 0.005$) and the three-way interaction approached significance ($F(2,48) = 2.76, p = 0.073$). Frequency by imageability interactions were significant when each group was examined individually (SD: $F(2,24) = 4.34, p < 0.05$; SA: $F(2,24) = 4.44, p < 0.05$). In the SA group, there was no frequency effect for low imageability ($t(12) = 1.03$) or medium imageability words ($t(12) = 1.08$) but for the highest imageability words, patients were more accurate for *low* frequency items ($t(12) = 3.59, p < 0.005$). In contrast, SD patients were more accurate for HF words at all levels of imageability ($t(12) > 2.29, p < 0.05$).

-Figure 3 around here-

To understand how word frequency, imageability and semantic diversity contributed to accuracy in each patient group we conducted a multiple regression analysis. For each word in the synonym judgment task, we first computed the proportion of patients in each group who performed correctly, providing a measure of how “difficult” comprehension of each item was for each patient group. We then began by examining, across items, the raw correlations between this measure and log word frequency, imageability, and semantic diversity in each group. As indicated by the correlation matrix in Table 5, comprehension in SD was correlated with frequency and imageability, while in SA imageability and semantic diversity were related to performance but frequency was not. Table 5 also shows, however, that semantic diversity is reliably correlated with both log word frequency and imageability.

To better understand the pattern of relationships, we conducted a multiple regression predicting itemwise comprehension from a linear combination of log word frequency, imageability and semantic diversity. Frequency and imageability were entered in the first step and diversity added as an additional predictor on the second step.⁴ This analysis was conducted separately for each patient group.

The first step of the regression confirmed the earlier analyses: comprehension in both groups was reliably predicted by imageability, whereas frequency predicted performance in the SD group only (see Table 6). When diversity was added as a predictor, the picture in the SD group was largely unchanged: inclusion of diversity did not significantly improve the fit of the regression model ($F(1,88) = 3.11, p = 0.08$). Effects of frequency and imageability remained significant while diversity was not a significant predictor. In the SA group the model that included semantic diversity accounted for an additional 11% of the variance in comprehension scores ($F(1,88) = 15.1, p < 0.001$). Moreover, diversity was the most important predictor of performance in this group, with beta weights indicating that more semantically diverse words were associated with poorer performance. Imageability remained a significant predictor and, most importantly, when diversity was included in the regression model a significant effect of frequency emerged. This was in the same direction as observed in healthy individuals and in SD patients: higher frequency predicted better comprehension. Frequency, however, remained the weakest predictor in this group and was a much less potent predictor than in the SD group.

-Tables 5 and 6 around here-

Discussion

In a verbal comprehension test, we found different frequency effects in SD and SA patients who were matched for overall performance. While HF words were much better comprehended by SD patients, there was no raw frequency effect in the SA group. Regression analyses showed that

⁴ Three probes were excluded from this analysis because semantic diversity values could not be calculated for them. One additional probe (*suffix*) was excluded because visual inspection of the relationship between diversity and SA accuracy indicated that it was an outlier. This item is also often failed by healthy participants.

comprehension in SD was influenced by both frequency and imageability but in SA only imageability predicted performance. Inclusion of semantic diversity in the regression model, however, changed the picture of SA considerably. Semantic diversity was the best predictor of performance in this group, in line with the view that semantic control deficits in these patients caused particular problems in judging the meanings of semantically diverse words. Moreover, when the influence of diversity was taken into account a significant positive frequency effect was revealed. The analysis suggests that the positive effects of word frequency are “masked” in SA as a result of (i) the negative effects of semantic diversity in this syndrome and (ii) the positive correlation between word frequency and semantic diversity. In SD, diversity was not a significant predictor of comprehension; instead, word frequency was the best predictor of accuracy, consistent with the hypothesis that HF words receive more robust representations within the semantic knowledge store.

General Discussion

Divergent effects of word frequency have been found in different patient groups with comprehension disorders. While SD patients show large frequency effects, in line with the HF word processing advantage found universally in healthy subjects, frequency effects tend to be small, absent, or even reversed in stroke aphasia. To address this issue, we explored comprehension in patients with damage to either the control or representational elements of semantic cognition, employing a novel measure of semantic diversity that derived estimates of contextual variability using latent semantic analysis. We predicted that HF words would be associated with higher semantic diversity and that high semantic diversity would be particularly detrimental to patients with semantic aphasia following stroke (SA) given their deficits in semantic control (Jefferies & Lambon Ralph, 2006). Taken together, these predictions provide a hypothesis explaining why word frequency has so little effect on semantic processing in SA: the advantages accrued by HF words are offset by the higher demands that semantically-diverse words place on the impaired control processes. In line with this hypothesis, we found (i) that HF words do indeed tend to appear in a more diverse set of contexts than LF words, (ii) that semantic diversity strongly predicted verbal comprehension in SA, and (iii) when this influence

was accounted for, a normal frequency effect emerged in SA. In contrast, semantic diversity had no influence on the performance of SD patients, who showed robust frequency effects. Both groups also showed better comprehension of high imageability words.

The strong effect of semantic diversity on performance in SA is consistent with a deficit in control and regulation of semantic knowledge. Words that appear in a diverse set of contexts develop a larger and more varied set of semantic associations, many of which will be irrelevant in any specific situation. Previous studies have shown that patients with SA have difficulty inhibiting irrelevant semantic associations in verbal and non-verbal domains (Corbett, Jefferies, & Lambon Ralph, 2009; Jefferies et al., 2007; Jefferies & Lambon Ralph, 2006; Noonan et al., 2010). In another synonym matching task, for example, patients found it difficult to choose the appropriate word when one of the foils shared a strong association with the probe (e.g., matching *piece* with *slice* when one of the foils was *cake*). The degree of constraint provided by the task is also a key determinant of success in these patients: they are better at tasks where the stimulus strongly implies a particular response. Synonym judgement provides only weak constraint over semantic activation, as the linguistic and environmental context, that would usually accompany a word and direct the semantic system towards the appropriate information, is absent. Instead, patients must rely on their faulty internal control processes to select the relevant aspects of the word's semantic representation. This process is less likely to be successful for high diversity words that have more complex and context-dependent representations.

Semantic diversity did not influence comprehension in patients with SD, consistent with the view that the semantic impairment in this disorder arises from damage to the semantic knowledge store itself (Jefferies & Lambon Ralph, 2006; Patterson et al., 2007; Rogers et al., 2004; Warrington, 1975). Instead, the strongest effect in this group was their sensitivity to word frequency, a phenomenon that has been demonstrated previously across a range of tasks (e.g., Funnell, 1995; Jefferies et al., 2009; Lambon Ralph et al., 1998; Patterson et al., 2006). The semantic representations of highly familiar words and concepts are more strongly instantiated because they are encountered more often and the semantic system has more opportunities to code their meanings. They also tend to be acquired earlier in life and to be encountered more often during the disease process, which may

contribute to their relative preservation in the face of representational breakdown (see Lambon Ralph et al., 1998).

Once the influence of semantic diversity had been accounted for, SA patients also showed better comprehension of HF words, albeit to a much lesser extent than in SD. HF words also show a processing advantage in healthy individuals, so this pattern likely reflects the amplification of an existing effect as a result of brain damage. That is, the frequency effect in accuracy in SA likely arises for the same reasons that such effects are observed in response time in healthy participants: more frequent words have more robust representations within the cortical semantic network and so are less affected by noisy or degraded processing in the system. This effect has not been observed previously because it has been counteracted by the higher semantic diversity of HF words, giving the false impression that frequency has no effect on this patient group.

Relationship to other work

Refractory access disorders. Lack of a frequency effect is also a feature of patients who present with “refractory access” deficits (Warrington & Cipolotti, 1996; Warrington & McCarthy, 1983; Warrington & Shallice, 1979). These individuals have a particular pattern of comprehension deficit characterised by decreases in accuracy when stimuli are repeated, when they are semantically related to one another and when they are presented at a fast rate. Patients usually also show response inconsistency, responding correctly to different stimuli on different occasions. This pattern is thought to reflect an increase in “refractoriness” during semantic access. Once a particular representation has been activated, it cannot be efficiently accessed again for a given refractory period. Gotts and Plaut (2002) have taken this proposal further by linking refractoriness to neuromodulatory systems that regulate neural firing. They have suggested refractory patients have a deficit in acetylcholine, which is needed to counteract synaptic depression processes that suppress the activity of neurons for a short time after they fire. In the absence of this neurotransmitter synaptic depression proceeds unchecked, preventing the necessary neural activity when the same stimulus is presented multiple times over a short period. On this view, frequency effects do not occur in these patients because their stronger

neural representations are subject to higher levels of synaptic depression, counteracting the expected processing advantage. Gotts and Plaut's theory is supported by a connectionist model in which damage to semantic representations yields a large frequency effect whilst simulated disruption to neuromodulatory processes results in very small frequency effects.

How do refractory access patients and the neuromodulatory account of their deficits relate to the SA patients reported in the present study? SA patients do show some of the features associated with refractory access disorders (Jefferies et al., 2007) but there are some notable differences. Refractory access patients typically have extremely severe comprehension impairments—they are almost always globally aphasic—while semantic control deficits are observed across a wide range of aphasia severity and the lack of a frequency effect is present across the severity spectrum (Jefferies et al., 2007). Only the most severe SA patients would qualify as 'classic' refractory patients. In addition, in Gotts and Plaut's (2002) model, the complete absence of a frequency effect was only observed with the most extreme simulated disruption of acetylcholine. It is possible, then, that there are qualitatively different ways that semantic processing may be disrupted without producing an effect of frequency, and that these different forms of disruption may lead to qualitatively different patterns of semantic impairment. A goal for future research is to understand the relationship between SA and refractory access disorders.

Effects of imageability. In addition to observing frequency effects in comprehension, we also found effects of imageability in both groups. These were reduced in magnitude in the SA group when diversity was considered but nevertheless continued to make a significant contribution—suggesting that such effects do not solely reflect the lesser degree of semantic diversity among more concrete concepts. Better processing of concrete words is commonly found in healthy individuals and in stroke aphasia generally (Coltheart, 1980; Degroot, 1989; Franklin, 1989; James, 1975; Katz & Goodglass, 1990; Kroll & Merves, 1986). In SD, the picture is more complex. There have been a number of case reports of patients who showed striking reverse imageability effects (Breedin, Saffran, & Coslett, 1994; Cipolotti & Warrington, 1995; Macoir, 2009; Papagno, Capasso, & Miceli, 2009; Warrington, 1975), although such effects are not consistently replicated in larger groups of unselected SD patients

and do not appear to be a typical feature of the disorder (Hoffman & Lambon Ralph, submitted; Jefferies et al., 2009). The positive imageability effect observed in both groups here has a number of possible explanations. First, abstract words may be more difficult to process because they depend on context to a greater degree to determine their meaning and this information is typically absent in experimental situations (Schwanenflugel, Harnishfeger, & Stowe, 1988; Schwanenflugel & Shoben, 1983). This explanation is supported by our finding that abstract words tended to appear in more diverse contexts. Even after accounting for this difference, however, an imageability effect remained in both groups, suggesting that their dependence on context is not the sole explanation of the processing disadvantage for abstract words.

Concrete words are also thought to have richer semantic representations because they are associated with sensory and motor information that is not available for abstract words (Paivio, 1986; Plaut & Shallice, 1993). As a consequence, degradation of semantic representations in SD might affect concrete concepts to a lesser extent. One might have expected the richer representations associated with concrete words to cause greater problems for the SA group, since richer representations by definition imply a greater amount of associated information that in turn might exert greater demands on cognition control. In this regard, it is important to distinguish *semantic richness* from *semantic diversity*. A concept that is semantically rich may not exert strong demands on cognitive control if the same sensory and motor information is always associated with a concept irrespective of context. A concrete word like *apple* is associated with a number of sensory features, such as shape, colour, texture and taste, but these properties are experienced fairly consistently whenever an example of the concept is encountered, so there is little need for the semantic control system to select or bias processing toward some of these features and not others. As suggested by the literature on cognitive control more generally, semantic control processes are likely to be required in cases where there are many *competing* representations, some of which must be potentiated and others inhibited. If the task permits simultaneous activation of many mutually consistent representations—like the various different sensory and motor properties of the apple—then there is little need to draw

on semantic control processes. Our particular comprehension task did not require patients to focus on specific aspects of a concept's sensory features as global similarity judgements were used.

Conclusion

The present study offers three major contributions to the study of human semantic cognition. First, we have introduced a method for measuring a word's semantic diversity—the extent to which it is encountered in contexts with quite different overall meanings. Second, we have shown that semantic diversity is associated with word frequency and imageability, with HF and less imageable words tending to be more diverse. Third, we have shown that the impact of these factors on impaired semantic processing differs qualitatively depending upon the underlying nature of the impairment. When the semantic impairment is caused by damage to the semantic network itself, the stronger representation of HF words within this network leads to better performance. When the impairment is caused by damage to semantic control systems, the processing advantage typically accruing to HF words is offset by the greater reliance semantically diverse words place on control processes—so that, when diversity is not taken into account, no frequency effect is observed. When the semantically diverse nature of HF words is statistically controlled, however, patients with semantic control deficits, like healthy individuals and patients with SD, show better performance for HF than LF words.

References

- Adelman, J. S., Brown, G. D. A., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science, 17*(9), 814-823.
- Baayen, R. H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM)*. Philadelphia, PA: Linguistic Data Consortium, University of Pennsylvania.
- Badre, D., & Wagner, A. D. (2002). Semantic retrieval, mnemonic control, and prefrontal cortex. *Behavioral and Cognitive Neuroscience Reviews, 1*, 206-218.
- Balota, D. A., & Chumbley, J. I. (1984). Are Lexical Decisions a Good Measure of Lexical Access - the Role of Word-Frequency in the Neglected Decision Stage. *Journal of Experimental Psychology-Human Perception and Performance, 10*(3), 340-357.
- Bozeat, S., Lambon Ralph, M. A., Patterson, K., Garrard, P., & Hodges, J. R. (2000). Non-verbal semantic impairment in semantic dementia. *Neuropsychologia, 38*(9), 1207-1215.
- Breedin, S. D., Saffran, E. M., & Coslett, H. B. (1994). Reversal of the concreteness effect in a patient with semantic dementia. *Cognitive Neuropsychology, 11*(6), 617-660.
- British National Corpus Consortium. (2007). *British National Corpus version 3 (BNC XML edition)*. Oxford: Oxford University Computing Services.
- Buchanan, L., Westbury, C., & Burgess, C. (2001). Characterizing semantic space: Neighborhood effects in word recognition. *Psychonomic Bulletin & Review, 8*(3), 531-544.
- Bullinaria, J. A., & Levy, J. P. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods, 39*(3), 510-526.
- Cipolotti, L., & Warrington, E. K. (1995). Semantic memory and reading abilities: A case report. *Journal of the International Neuropsychological Society, 1*, 104-110.
- Coltheart, M. (1980). Deep dyslexia: A review of the syndrome. In M. Coltheart, K. Patterson & J. C. Marshall (Eds.), *Deep Dyslexia*. London: Routledge and Kegan Paul.
- Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 33*(Nov), 497-505.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108*(1), 204-256.
- Corbett, F., Jefferies, E., Ehsan, S., & Lambon Ralph, M. A. (2009). Different impairments of semantic cognition in semantic dementia and semantic aphasia: Evidence from the non-verbal domain. *Brain, 132*, 2593-2608.
- Corbett, F., Jefferies, E., & Lambon Ralph, M. A. (2009). Exploring multimodal semantic control impairments in semantic aphasia: Evidence from naturalistic object use. *Neuropsychologia, 47*, 2721-2731.
- Crutch, S. J., & Warrington, E. K. (2005). Abstract and concrete concepts have structurally different representational frameworks. *Brain, 128*, 615-627.
- Degroot, A. M. B. (1989). Representational aspects of word imageability and word frequency as assessed through word association. *Journal of Experimental Psychology: Learning Memory and Cognition, 15*(5), 824-845.
- Dell, G. S. (1989). The retrieval of phonological forms in production: Tests of predictions from a connectionist model. In W. Marslen Wilson (Ed.), *Lexical representation and process*. Cambridge, MA: MIT Press.
- Durda, K., & Buchanan, L. (2008). WINDSORS: Windsor improved norms of distance and similarity of representations of semantics. *Behavior Research Methods, 40*(3), 705-712.
- Firth, J. R. (1957). A synopsis of linguistic theory. In *Studies in linguistic analysis*. Oxford: Blackwell.
- Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior, 12*, 627-635.
- Franklin, S. (1989). Dissociations in auditory word comprehension; Evidence from nine fluent aphasic patients. *Aphasiology, 3*, 189-207.
- Funnell, E. (1995). Objects and properties: A study of the breakdown of semantic memory. *Memory, 3*(3-4), 497-518.

- Garavan, H., Ross, T. J., Li, S. J., & Stein, E. A. (2000). A parametric manipulation of central executive functioning. *Cerebral Cortex*, *10*(6), 585-592.
- Garrard, P., & Carroll, E. (2006). Lost in semantic space: a multi-modal, non-verbal assessment of feature knowledge in semantic dementia. *Brain*, *129*, 1152-1163.
- Giles, J. T., Wo, L., & Berry, M. W. (2002). GTP (General Text Parser) software for text mining. In H. Bozdogan (Ed.), *Statistical data mining and knowledge*. Boca Raton: CRC Press.
- Gotts, S. J., & Plaut, D. C. (2002). The impact of synaptic depression following brain damage: A connectionist account of "access/refractory" and "degraded-store" semantic impairments. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(3), 187-213.
- Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, *114*(2), 211-244.
- Harris, Z. (1954). Distributional structure. *Word*, *10*, 146-162.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*(5), 393-402.
- Hoffman, P., Jefferies, E., & Lambon Ralph, M. A. (in press). Ventrolateral prefrontal cortex plays an executive regulation role in comprehension of abstract words: Convergent neuropsychological and rTMS evidence. *Journal of Neuroscience*.
- Hoffman, P., Jefferies, E., & Lambon Ralph, M. A. (submitted). Remembering 'zeal' but not 'thing': Reverse frequency effects as a consequence of deregulated semantic processing.
- Hoffman, P., & Lambon Ralph, M. A. (submitted). Reverse concreteness effects are not a typical feature of semantic dementia: Evidence for the hub-and-spoke model of conceptual representation.
- Hon, N., Epstein, R. A., Owen, A. M., & Duncan, J. (2006). Frontoparietal activity with minimal decision and control. *Journal of Neuroscience*, *26*(38), 9805-9809.
- James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *10*(4), 130-136.
- Jefferies, E., Baker, S. S., Doran, M., & Lambon Ralph, M. A. (2007). Refractory effects in stroke aphasia: A consequence of poor semantic control. *Neuropsychologia*, *45*, 1065-1079.
- Jefferies, E., & Lambon Ralph, M. A. (2006). Semantic impairment in stroke aphasia vs. semantic dementia: A case-series comparison. *Brain*, *129*, 2132-2147.
- 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.
- Jefferies, E., Patterson, K., & Lambon Ralph, M. A. (2008). Deficits of knowledge versus executive control in semantic cognition: Insights from cued naming. *Neuropsychologia*, *46*, 649-658.
- Katz, R. B., & Goodglass, H. (1990). Deep dysphasia: Analysis of a rare form of repetition disorder. *Brain and Language*, *39*(1), 153-185.
- Koenig, K., & Grossman, M. (2007). Process and content in semantic memory. In J. Hart & M. A. Kraut (Eds.), *Neural basis of semantic memory*. Cambridge: Cambridge University Press.
- Kroll, J. F., & Merves, J. S. (1986). Lexical access for concrete and abstract words. *Journal of Experimental Psychology: Learning Memory and Cognition*, *12*, 92-107.
- Lambon Ralph, M. A., Graham, K. S., Ellis, A. W., & Hodges, J. R. (1998). Naming in semantic dementia - what matters? *Neuropsychologia*, *36*(8), 775-784.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction and representation of knowledge. *Psychological Review*, *104*, 211-240.
- Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods Instruments & Computers*, *28*(2), 203-208.
- Macoir, J. (2009). Is a plum a memory problem? Longitudinal study of the reversal of the concreteness effect in a patient with semantic dementia. *Neuropsychologia*, *47*, 518-535.
- Marshall, J., Pring, T., Chiat, S., & Robson, J. (2001). When ottoman is easier than chair: An inverse frequency effect in jargon aphasia. *Cortex*, *37*(1), 33-53.
- Martin, A. J. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, *58*, 25-45.

- Mirman, D., & Magnuson, J. S. (2008). Attractor dynamics and semantic neighborhood density: Processing is slowed by near neighbors and speeded by distant neighbors. *Journal of Experimental Psychology-Learning Memory and Cognition*, *34*(1), 65-79.
- Nagel, I. E., Schumacher, E. H., Goebel, R., & D'Esposito, M. (2008). Functional MRI investigation of verbal selection mechanisms in lateral prefrontal cortex. *Neuroimage*, *43*(4), 801-807.
- Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S., et al. (1998). Frontotemporal lobar degeneration: A consensus on clinical diagnostic criteria. *Neurology*, *51*(6), 1546-1554.
- Nelson, D. L., & McEvoy, C. L. (2000). What is this thing called frequency? *Memory & Cognition*, *28*(4), 509-522.
- Nickels, L., & Howard, D. (1995). Aphasic naming: What matters? *Neuropsychologia*, *33*(10), 1281-1303.
- Noonan, K. A., Jefferies, E., Corbett, F., & Lambon Ralph, M. A. (2010). Elucidating the nature of deregulated semantic cognition in semantic aphasia: Evidence for the roles of the prefrontal and temporoparietal cortices. *Journal of Cognitive Neuroscience*, *22*, 1597-1613.
- Noonan, K. A., Jefferies, E., Visser, M., & Lambon Ralph, M. A. (submitted). Aligning evidence from functional neuroimaging and neuropsychology for the neural network underpinning semantic control: A meta-analytic investigation.
- Norris, D. (1994). A quantitative multiple-levels model of reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 1212-1232.
- Oldfield, R. C. (1966). Things, words and the brain. *Quarterly Journal of Experimental Psychology*, *18*, 3-16.
- Paivio, A. (1986). *Mental representations: A dual-coding approach*. Oxford: Oxford University Press.
- Papagno, C., Capasso, R., & Miceli, G. (2009). Reversed concreteness effect for nouns in a subject with semantic dementia. *Neuropsychologia*, *47*, 1138-1148.
- Patterson, K., Lambon Ralph, M. A., Jefferies, E., Woollams, A., Jones, R., Hodges, J. R., et al. (2006). "Presemantic" cognition in semantic dementia: Six deficits in search of an explanation. *Journal of Cognitive Neuroscience*, *18*(2), 169-183.
- 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*(12), 976-987.
- Peers, P. V., Ludwig, C. J. H., Rorden, C., Cusack, R., Bonfiglioli, C., Bundesen, C., et al. (2005). Attentional functions of parietal and frontal cortex. *Cerebral Cortex*, *15*, 1469-1484.
- Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes*, *12*, 765-805.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*(1), 56-115.
- Plaut, D. C., & Shallice, T. (1993). Deep dyslexia: A case study in connectionist neuropsychology. *Cognitive Neuropsychology*, *10*, 377-500.
- Pobric, G., Jefferies, E., & Lambon Ralph, M. A. (2010). Category-specific vs. category-general semantic impairment induced by transcranial magnetic stimulation. *Current Biology*, *20*, 964-968.
- Raven, J. C. (1962). Coloured progressive matrices sets A, AB, B. London: H. K. Lewis.
- Recchia, G., & Jones, M. N. (2009). More data trumps smarter algorithms: Comparing pointwise mutual information with latent semantic analysis. *Behavior Research Methods*, *41*(3), 647-656.
- Reder, L. M., Anderson, J. R., & Bjork, R. A. (1974). Semantic interpretation of encoding specificity. *Journal of Experimental Psychology*, *102*, 648-656.
- Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex*, *15*(8), 1261-1269.
- Rodd, J. M., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, *46*(2), 245-266.

- Rogers, T. T., Lambon Ralph, M. A., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., et al. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, *111*(1), 205-235.
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.
- Saffran, E. M., Bogyo, L., Schwartz, M. F., & Marin, O. S. M. (1980). Does deep dyslexia reflect right hemisphere reading? In M. Coltheart, J. C. Marshall & K. E. Patterson (Eds.), *Deep dyslexia* (pp. 361-406). London: Routledge & Kegan Paul.
- Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context Availability and Lexical Decisions for Abstract and Concrete Words. *Journal of Memory and Language*, *27*(5), 499-520.
- Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential Context Effects in the Comprehension of Abstract and Concrete Verbal Materials. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *9*(1), 82-102.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523-568.
- Shaoul, C., & Westbury, C. (2006). Word frequency effects in high-dimensional co-occurrence models: A new approach. *Behavior Research Methods*, *38*(2), 190-195.
- Simmons, W. K., & Martin, A. (2009). The anterior temporal lobes and the functional architecture of semantic memory. *Journal of the International Neuropsychological Society*, *15*(5), 645-649.
- Snowden, J. S., Goulding, P. J., & Neary, D. (1989). Semantic dementia: a form of circumscribed cerebral atrophy. *Behavioural Neurology*, *2*, 167-182.
- Stemberger, J. P. (1985). An interactive activation model of language production. In A. W. Ellis (Ed.), *Progress in the psychology of language production* (Vol. 1). London: Lawrence Erlbaum Associates Ltd.
- Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America*, *94*, 14792-14797.
- Visser, M., Jefferies, E., & Lambon Ralph, M. A. (2010). Semantic processing in the anterior temporal lobes: A meta-analysis of the functional neuroimaging literature. *Journal of Cognitive Neuroscience*, *22*, 1083-1094.
- Wagner, A. D., Pare-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron*, *31*(2), 329-338.
- Warrington, E. K. (1975). The selective impairment of semantic memory. *Quarterly Journal of Experimental Psychology*, *27*, 635-657.
- Warrington, E. K., & Cipolotti, L. (1996). Word comprehension: The distinction between refractory and storage impairments. *Brain*, *119*, 611-625.
- Warrington, E. K., & James, M. (1991). *The Visual Object and Space Perception battery*. Bury St. Edmunds, Suffolk: Thames Valley Test Company.
- Warrington, E. K., & McCarthy, R. A. (1983). Category specific access dysphasia. *Brain*, *106*, 859-878.
- Warrington, E. K., & Shallice, T. (1979). Semantic access dyslexia. *Brain*, *102*(MAR), 43-63.
- Wechsler, D. (1987). *Wechsler memory scale: Revised (WMS-R)*. New York: Psychological Corporation.
- Welbourne, S. R., & Lambon Ralph, M. A. (2005). Exploring the impact of plasticity-related recovery after brain damage in a connectionist model of single-word reading. *Cognitive Affective & Behavioral Neuroscience*, *5*(1), 77-92.
- Wittgenstein, L. (1953). *Philosophical investigations*. Oxford: Blackwell.
- Woollams, A., Cooper-Pye, E., Hodges, J. R., & Patterson, K. (2008). Anomia: A doubly typical signature of semantic dementia. *Neuropsychologia*, *46*, 2503-2514.
- Zemleni, M. Z., Renken, R., Hoeks, J. C. J., Hoogduin, J. M., & Stowe, L. A. (2007). Semantic ambiguity processing in sentence context: Evidence from event-related fMRI. *Neuroimage*, *34*(3), 1270-1279.

Table 1: Examples of semantic diversity values

Ten lowest values		Ten highest values			
Word	SemD	Word	SemD	Word	SemD
verdant	0.30	harpsichord	0.63	one	2.36
caddie	0.30	sonic	0.80	part	2.35
catfish	0.36	yolk	1.04	have	2.35
gastric	0.38	gallery	1.23	may	2.35
duvet	0.46	traitor	1.45	has	2.34
polyp	0.49	muddy	1.60	new	2.34
sonata	0.49	suppression	1.76	being	2.34
melee	0.49	big	1.95	following	2.33
adverb	0.50	life	2.13	used	2.32
fielder	0.52	similar	2.30	number	2.32

The ten lowest and ten highest values obtained are shown, along with a representative selection of other words. SemD = semantic diversity.

Table 2: Correlation matrix for lexical and semantic variables

Variable	Log frequency	Imageability	Log number of senses
Semantic diversity	.534***	-.344***	.412***
Log frequency	--	-.044**	.493***
Imageability		--	.027
Log number of senses			--

** = $p < 0.01$; *** = $p < 0.001$.

Table 3: Results of simultaneous regression analysis on semantic diversity values

	<i>B</i>	Standard error	β
Log frequency	.202	.006	.404***
Imageability	-.001	.000003	-.341***
Log number of senses	.193	.011	.222***

*** = $p < 0.001$.

Table 4: Demographic and neuropsychological data for patients

Test	SA		SD		Healthy controls	
	Mean	<i>s.d.</i>	Mean	<i>s.d.</i>	Mean	<i>s.d.</i>
Age	66.2	12.5	63.4	7.2	-	-
Education (leaving age)	15.6	1.7	16.0	2.6	-	-
Synonym judgement test /96	68.7	10.5	69.3	10.8	94.5	1.8
<i>Cambridge Semantic Battery^a</i>						
Picture naming /64	28.2	21.7	34.3	17.7	62.3	1.6
Word-picture matching /64	51.2	10.3	51.7	14.4	63.8	1.4
Camels and Cactus test: words /64	39.3	10.2	42.3	12.2	60.7	2.1
Camels and Cactus test: pictures /64	38.2	11.9	44.2	12.1	59.0	3.1
Category fluency (8 categories)	25.5	15.6	42.3	19.3	113.9	12.3
<i>VOSP^b</i>						
Dot counting /10	9.3	4.1	10.0	0	9.9	0.2
Position discrimination /20	17.4	3.0	19.5	1.4	19.6	0.9
Number location /10	8.8	3.9	9.5	0.9	9.4	1.1
Raven's coloured progressive matrices /36	19.5*	6.8	33.2	5.4	-	-
Digit span forwards	4.7*	1.2	7.2	1.1	6.8	0.9
Digit span backwards	1.8*	0.8	5.4	1.1	4.7	1.2

* indicates SA mean is significantly lower than SD (two-tailed t-test, $p < 0.05$). ^a Bozeat et al. (2000).

^b Visual Object and Space Perception battery (Warrington & James, 1991).

Table 5: Correlation matrix for variables in synonym judgement analysis

Variable	SD accuracy	Log frequency	Imageability	Semantic diversity
1. SA accuracy	.412***	-.053	.465***	-.438***
2. SD accuracy	--	.528***	.413***	.069
3. Log frequency		--	-.021	.596***
4. Imageability			--	-.361***
5. Semantic diversity				--

*** = $p < 0.001$.

Table 6: Results of regression analyses on synonym judgement accuracy

Predictor	SA group				SD group			
	R^2	B	Standard error	β	R^2	B	Standard error	β
<i>Step 1</i>	.219***				.460***			
Log frequency		-.013	.028	-.043		.173	.025	.537***
Imageability		.001	.0001	.464***		.001	.0001	.425***
<i>Step 2</i>	.333***				.478***			
Log frequency		.068	.033	.231*		.209	.032	.648***
Imageability		.0004	.0001	.302**		.001	.0001	.360***
Semantic diversity		-.286	.074	-.467***		-.125	.071	-.188

* = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$.

Figure 1: Relationship between semantic diversity and other lexical-semantic variables

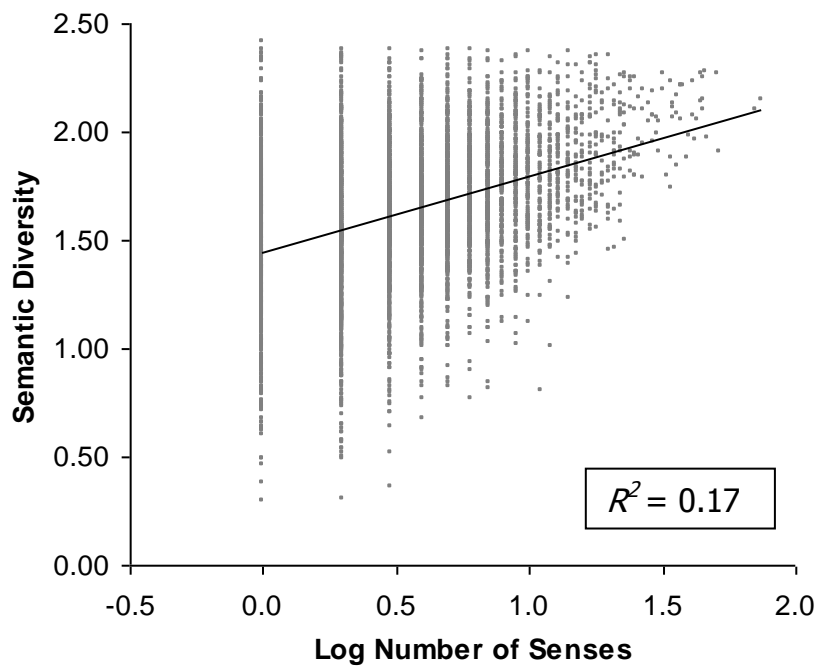
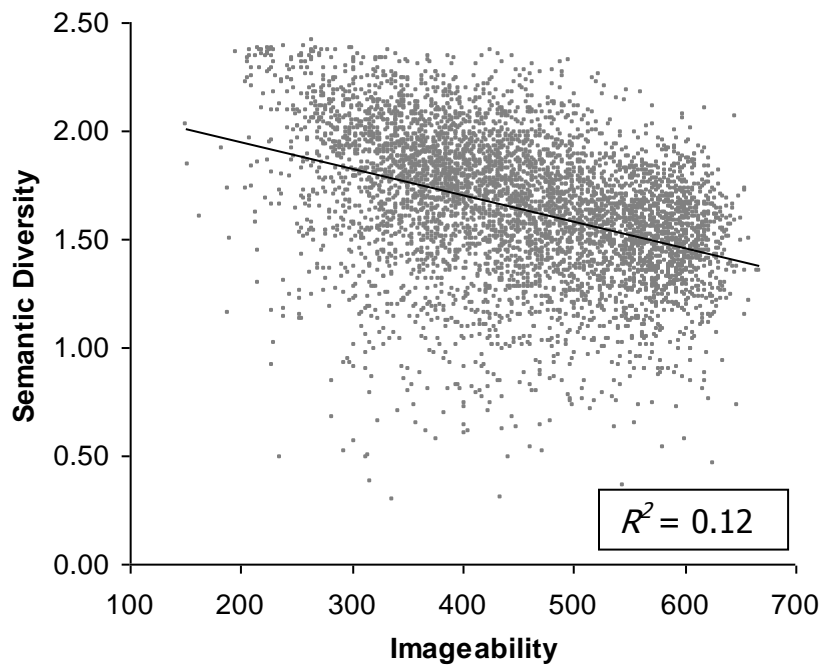
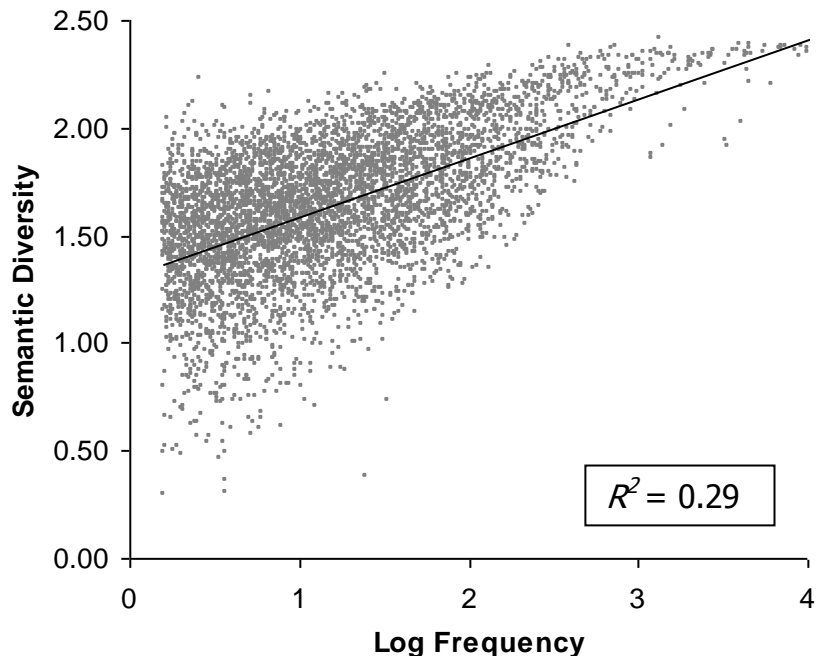


Figure 2: Mean semantic diversity values plotted in frequency x imageability space

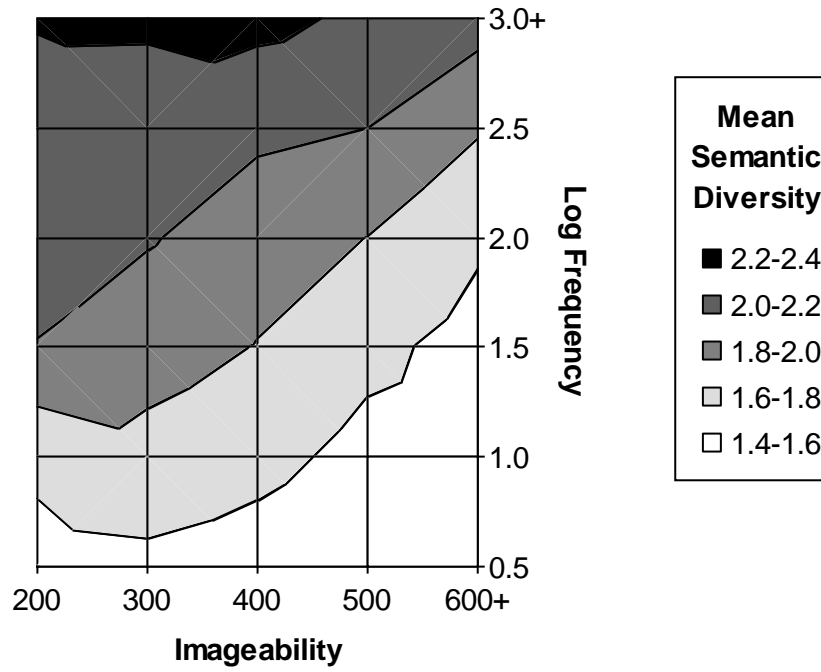


Figure 3: Synonym judgement performance in semantic dementia and semantic aphasia

