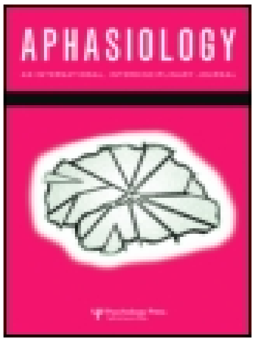
Novel matched stimuli for assessment of lexical semantics

DYSON, Lucy, MORGAN, Jane <http://orcid.org/0000-0001-8024-028X> and HERBERT, Ruth

Published version

DYSON, Lucy, MORGAN, Jane and HERBERT, Ruth (2021). Novel matched stimuli for assessment of lexical semantics. Aphasiology.



Aphasiology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/paph20

Novel matched stimuli for assessment of lexical semantics

Lucy Dyson, Jane Morgan & Ruth Herbert

APHASIOLOGY
https://doi.org/10.1080/02687038.2021.1924354

Routledge
Taylor & Francis Group

ARTICLE

OPEN ACCESS Check for updates

Novel matched stimuli for assessment of lexical semantics

Lucy Dyson [a], Jane Morgan [b] and Ruth Herbert [a]

[a]Division of Human Communication Sciences, Health Sciences School, University of Sheffield, Sheffield, UK; [b]Department of Psychology, Sheffield Hallam University, Sheffield, UK

ABSTRACT

Background: Diagnosis of semantic impairment in stroke and progressive neuro-cognitive conditions is typically facilitated using tests of word comprehension, such as word-picture matching. Many of these tests are not controlled for psycholinguistic variables or the semantic relationships between competitor stimuli and involve pictures which are not controlled for ease of access. Semantic assessment also demands additional cognitive resources, such as explicit decision-making and suppression of semantic competitors. These factors may all confound test performance and subsequent diagnosis.

Aims: To develop novel semantic test stimuli for three new semantic processing assessments, which are controlled for psycholinguistic variables, semantic relationship between stimuli, and visual similarity between images presented simultaneously. An additional aim included matching stimuli for these variables across three tests: semantic priming, word-picture verification, and word-picture matching, to allow direct comparison of performance on tests that differ in terms of the additional cognitive demands involved, with priming entailing implicit semantic processing.

Methods & Procedures: In phase one, novel stimuli were developed for the three semantic processing tests. Existing databases were searched for values to match stimuli psycholinguistic variables. In phase two, new data were collected from control participants regarding the semantic and visual similarity of stimuli presented simultaneously.

Outcomes & Results: Data for three sets of target and distractor stimuli are presented, which are psycholinguistically matched within and between the three semantic tests for concreteness, imageability, age of acquisition, frequency, word length, and emotional valence. "Semantic" relationships between pairs of stimuli are differentiated by semantic similarity (dog-cat) or association (dog-lead). Visual similarity is controlled between images presented in an array.

Conclusions: The data provided ensure that test performance across three semantic tasks, differing in additional cognitive demands, can be directly compared in people with potential semantic deficits. This is the first such study to provide control of stimuli within and across a range of semantic tests. Patterns of performance via test reaction time and accuracy data may reveal semantic deficit or competence, contributing to more accurate diagnosis and appropriate therapy choice.

ARTICLE HISTORY
Received 31 July 2020
Accepted 16 April 2021

KEYWORDS
Aphasia; semantics; association; assessment; semantic priming

CONTACT Lucy Dyson lucy.dyson@sheffield.ac.uk Division of Human Communication Sciences, Health Sciences School, University of Sheffield, 362 Mushroom Lane, Sheffield S10 2TS, UK

Supplemental data for this article can be accessed here.

Introduction

Neuropsychology and semantics

People with a range of neuro-pathologies may present with semantic processing difficulties. This includes aphasia post-stroke and progressive conditions notably semantic dementia and dementia of the Alzheimer's type; however, the nature of the semantic impairment differs between these groups. Aphasia involves lesions to perisylvian regions in the dominant hemisphere affecting frontal and temporoparietal regions primarily (e.g., Jefferies & Lambon Ralph, 2006). Semantic aphasia has been identified as a subtype of aphasia in which lexical semantic processing is disproportionately affected, due to difficulties with accessing semantic information (e.g., Warrington & Cipolotti, 1996). Different accounts of this type of deficit exist. Jefferies and Lambon Ralph (2006) proposed that impairment to control mechanisms executed by frontal lobe systems explains the type of semantic deficit found in semantic aphasia (see Corbett et al., 2009; Jefferies et al., 2007; Jefferies & Lambon Ralph, 2006; Noonan et al., 2010). Dell and colleagues in contrast developed explanations based on impairments to activation within their interactive activation theory of sentence processing which was first expounded in Dell et al. (1997). Various mechanisms within the model explain impaired semantic processing, including weakened activation, which negatively affects transmission between processing levels, and slower decay rates to activated nodes, which impacts on processing of subsequent stimuli due to maintained activation within non-target nodes (see Dell et al., 2007, 1997; Martin & Dell, 2019; Schwartz et al., 2006). In contrast, semantic dementia arises due to atrophy of the anterior temporal lobes (Mummery et al., 2000; Nestor et al., 2006) which have been identified as the seat of semantic knowledge (Lambon Ralph et al., 2017). Semantic dementia is characterised by a progressive, generalised degradation of stored semantic knowledge (Hodges et al., 1992; Mummery et al., 2000; Patterson et al., 2007), resulting in a multi-modality semantic impairment affecting lexical and non-verbal semantics, such as the ability to extract meaning from sounds, smell and touch (Bozeat et al., 2000).

Warrington and Shallice (1979) identified several behavioural phenomena that distinguish between semantic storage versus semantic access deficits, for example, cues are effective for access deficits but not storage deficits (Mirman & Britt, 2014). The theory ascribing the deficit in semantic aphasia to deficit in semantic control was proposed by Jefferies and Lambon Ralph (2006). Semantic control is defined as the precise activation of semantic features, inhibition of related items, and selection of the target through activation. The evidence in support of this view comes from a small number of participants with aphasia who perform more poorly on semantic tasks requiring higher executive control demands such as inferring the relevance of specific features of semantic meaning from a stimuli set (Jefferies & Lambon Ralph, 2006). The authors also found significant correlations between scores on semantic tests (word-picture matching, Pyramids and Palm Trees and picture naming) and an executive skill factor demonstrated in scores on Raven's Coloured Progressive Matrices test (Raven, 1956), and the Wisconsin Card Sorting Test (Heaton et al., 1993) in participants with semantic aphasia. This finding has been used to support claims of impaired executive control of semantics, related to more general executive control dysfunction (see also Corbett et al., 2009; Jefferies et al., 2007;

Noonan et al., 2010). More recent studies have failed to replicate a correlation between scores on semantic and executive function tests in people with aphasia however (Chapman et al., 2020; Dyson et al., 2021). As a range of semantic task performance data contribute to the diagnosis of semantic impairments, and related theoretical accounts, it is important to consider the potential difficulties in interpreting test data from semantic assessments.

Explicit versus implicit semantic processing tasks

Assessments used to diagnose semantic difficulties in aphasia and neuro-cognitive conditions often involve explicit matching of cross-modal stimuli, and deciding between closely related pairs of stimuli, in tasks such as word-picture matching and word-picture verification (e.g., Breese & Hillis, 2004; Howard & Gatehouse, 2006; Howard & Orchard-Lisle, 1984). Such tasks involve more than semantic processing however, requiring explicit decision-making, and resolution of semantic competition. In addition to semantic processing, task performance also draws on a range of cognitive skills that may also be impaired in the individual with aphasia or progressive neuro-cognitive condition. For example, impaired attention may affect language comprehension (Murray, 2002; Tabor Connor & Fucetola, 2011), and impaired visual-perceptual skills may impact on processing of written stimuli or the ability to visually scan and process images within a word-picture matching array (Heuer & Hallowell, 2007). It is therefore arguable whether these widely used tests are probing lexical semantic knowledge transparently, as additional cognitive demands and non-linguistic deficits may impact on performance (Ivanova & Hallowell, 2013).

Semantic priming provides a possible alternative or addition to the diagnostic process (Marinis, 2010; Milberg & Blumstein, 1981). Priming arguably probes more directly the real-time processing of language, by measuring participants' unconscious reactions to stimuli, i.e., enhanced speed of response or accuracy to a target if a similar stimulus, known as a prime, has been presented previously, compared to when the target is preceded by a unrelated stimulus (McNamara, 2005; McNamara & Holbrook, 2003; Meyer & Schvaneveldt, 1971; Tulving & Schacter, 1990). For example, the word *apple* would typically be responded to faster in a task if preceded by the word *pear* than if preceded by the word *guitar*. The participant is not consciously aware of the relationship between test stimuli and is not making explicit decisions about word meaning. This unconscious or implicit nature of processing in semantic priming contrasts with that of explicit tests such as word-picture matching. Semantic priming typically involves one stimulus at a time and lexical decision through a yes/no choice, and arguably therefore also places far fewer demands on executive functions.

The possibility that explicit tasks such as word-picture matching are harder for people than implicit tasks such as priming was demonstrated by Milberg and Blumstein (1981), who found retained semantic priming effects in people diagnosed with lexical semantic deficits in the context of Wernicke's aphasia. This evidence indicated that the perceived semantic deficits were an artefact of the explicit tasks. Subsequent studies with people with fluent aphasia and impaired performance on explicit tests of semantic processing revealed positive semantic priming effects, and this has been interpreted as evidence for retained semantic knowledge (Baum, 1997; Blumstein et al., 1982; Milberg et al., 1987; Yee, 2005). This evidence supports the claim that impaired performance in explicit tasks is not due to semantic damage or even to impaired semantic control but may be due to impairment of one or more supporting cognitive mechanisms, which are recruited by the complex demands of the assessment methods used. The concept of semantic access remains underspecified, but Mirman and Britt (2014) refer to selection, activation and inhibition of semantic representations. Positive semantic priming effects can be viewed as evidence of retained semantic activation, whilst impaired performance on explicit offline tasks may be due to difficulties with semantic selection and inhibition. Thus, comparing performance across explicit and implicit tasks may shed light on the nature of the difficulties an individual is facing, refine diagnosis more precisely, and hence lead to more appropriate interventions.

Test construction in the current study

In this study we aimed to develop matched stimuli for two explicit tasks, word-picture matching and word-picture verification, both involving explicit, offline semantic judgements and resolution of semantic competition, and one implicit online task, semantic priming. The resulting set of three tasks can be used to identify retained semantic activation through the semantic priming task, and to compare this evidence with participants' ability to handle selection and semantic competition in the explicit tasks. The tasks all depend on the semantic relationship between prime and target or target and distractor, so careful selection and control of these relationships was built into the stimuli sets. Many existing tests such as the PALPA (Kay et al., 1992) do not control fully for semantic relationship, psycholinguistic variables, and visual similarity between stimuli, and again these were all addressed in our test construction. These issues are discussed below.

Task artefacts

Semantic competition from test stimuli

In many semantic assessments the presence of competitor stimuli forms the critical component. For example, word-picture matching involves making an explicit decision to select one picture from an array, to match a spoken or written word. Distractor stimuli can be manipulated in terms of semantic, phonological or visual relationship to the target to ascertain the impact of damage to these processing routines. Several theorists have proposed that error patterns in lexical semantic tasks reflect the level of functional impairment in the language processing system (see Butterworth et al., 1984; Cutler, 1981; Kay et al., 1992, 1996; Shallice, 1987). For example, a preponderance of semantic selection errors may indicate impaired semantic processing (e.g., KE: Hillis et al., 1990; JCU: Howard & Orchard-Lisle, 1984; Howard & Gatehouse, 2006), or impaired access to semantics from the lexicon or access to the lexicon from semantics (e.g., AR: Warrington & Shallice, 1979). Word-picture verification involves the presentation of single word and picture stimuli, in pairs that are congruent where the target word is presented with the target picture, or incongruent where the semantic distractor word is presented with the target picture. Successful completion involves accepting congruent pairs and rejecting incongruent pairs through yes/no judgment (Breese & Hillis, 2004; Howard & Franklin, 1988; Howard & Gatehouse, 2006; Howard & Orchard-Lisle, 1984; Morris & Franklin, 2012;

Rapp & Caramazza, 2002). It is well recorded that people with aphasia show semantic interference in tasks involving semantically related stimuli, including in word picture matching tasks (Crutch & Warrington, 2005; Harvey & Schnur, 2015; Thompson et al., 2015; Warrington & Cipolotti, 1996; Warrington & McCarthy, 1983, 1987). It is therefore problematic that explicit semantic judgment tasks involve the presence of at least one semantic competitor and hence involve resolution of competition.

Relationships between word pairs

The nature of the relationship between the target and distractor is a further component which can affect processing in semantic tasks. The pairs may be semantically similar i.e., they share many semantic features, such as target *dog* paired with distractor *cat*, or they may be associated, such as target *dog* paired with distractor *lead*. Concepts which are thematically associated are those which co-occur frequently in situations or events, but which do not share semantic features (e.g., fork-spaghetti) (Lin & Murphy, 2001; Mirman & Graziano, 2012a; Moss et al., 1995). In existing word-picture matching tasks, the distinction between semantic similarity versus association between targets and distractors is either not made or adequately controlled. Cole-Virtue and Nickels (2004a) collected control participants' ratings of semantic relationships between targets and distractors from the PALPA spoken word-picture matching task, which revealed inconsistencies in the pairs, with roughly a quarter of test items having associative relationships as opposed to the claimed semantic similarity. Inconsistent patterns found in semantic priming effect data have been hypothesised to relate to the type of relationship between prime and target stimuli, with evidence of the separate contribution of both semantic similarity and association (for reviews see Hutchison, 2003; Lucas, 2000). Semantic similarity and association between stimuli pairs therefore need to be controlled in semantic tasks to ensure parity in task difficulty and processing requirements.

Psycholinguistic variables

A further factor influencing the assessment of semantics concerns the degree to which psycholinguistic variables are controlled in stimulus selection. Processing of lexical stimuli is affected by key factors, in particular imageability, lexical frequency, and word length. In word comprehension tasks, people with aphasia process words which have higher imageability or concreteness ratings more easily (Franklin, 1989; Franklin et al., 1995, 1996). The matching of target and distractor frequency is proposed to be important in the choice of word-picture matching stimuli, as items in an array with a higher frequency may possess a processing advantage to simultaneously presented items of lower frequency (Cole-Virtue & Nickels, 2004b). Word length has been shown to affect response speed and accuracy in people with aphasia, with shorter words processed more easily than longer words (e.g., Nickels & Howard, 1995). To ensure reliability of assessments the semantic word pairs need to be matched for key variables, for example, to ensure that high frequency targets are not paired with low frequency distractors, or vice versa.

Visual characteristics of test stimuli

Pictures used in semantic tests are not routinely controlled for ease of access, and this is a further variable that could impact on task performance. In some word-picture matching tests visually similar items are included in the array in an attempt to identify perceptual deficit, for example, in the PALPA (Kay et al., 1992). Inconsistencies in visual similarity between target items and visually unrelated distractors are reported however, for example, close semantic distractors are rated as more visually similar to the targets than the distant distractors (Cole-Virtue & Nickels, 2004a). It is also recognised within the literature that object recognition is facilitated by providing colour information (Rossion & Pourtois, 2001, 2004), which is not consistently provided in semantic assessments, with many using black and white line drawings. In some instances, test stimuli include a mix of line-drawings photographic images, with varying levels of visual definition (e.g., see Camel and Cactus Test, Bozeat et al., 2000) and concern has been expressed by clinicians and researchers regarding the quality of PALPA picture stimuli (black line drawings), with the use of colour and photographic images proposed to improve this confound (Bate et al., 2010). A preference for the use of photographs in materials over line drawings or symbols has also been reported by people with aphasia (Rose et al., 2012). As the visual similarity between test images and the accessibility of images can influence participant performance, this also needs to be controlled in the design of semantic assessments.

Rationale and aims

In this study we developed novel stimuli for three tasks: two explicit offline tasks, word-picture verification and word-picture matching, and one implicit online task, semantic priming, involving unconscious processing of semantic knowledge through lexical decision to single written words. Two explicit semantic tasks were included in order to allow a comparison of the impact of the number of distractor stimuli on performance, and within our word-picture verification task, to compare the congruent and the incongruent conditions which give a direct window onto the impact of semantic competition on processing. Previous studies of lexical semantic processing have used either explicit tasks (e.g., Howard & Gatehouse, 2006; Rapp & Caramazza, 2002), or implicit semantic priming tasks, with no study comparing these directly using controlled and matched stimuli. In the study described here we aimed to rectify this situation by designing matched sets of stimuli for three new lexical semantic processing tests. The tests are controlled for psycholinguistic variables, the semantic or associative relationship between stimuli, and the visual similarity between images presented simultaneously. In addition to control within the individual tests, stimuli are matched across the three tests, which was essential in order to allow comparisons in performance between tasks. The ultimate aim was to use the new stimuli to examine semantic processing in healthy participants and in people with aphasia, the results of which are reported in Dyson et al. (2021). The work described here focuses on the development and matching of the task stimuli.

The aims of this work were four-fold. To develop sets of prime-target stimuli (semantic priming task) and target-distractor stimuli (word-picture verification and word-picture matching tasks) matched for a range of psycholinguistic variables within and between sets. To ensure each set included a subset with semantically similar word pairs and a subset with associated word pairs, again matched across tasks. To collect normative data regarding semantic similarity for all word pairs within each task, and to match for this variable within and between tasks. To collect normative data regarding visual similarity for word pairs in the word-picture matching task. The resulting stimuli are aimed at researchers and clinicians working with populations with potential semantic processing impairment.

Methods

Overview of phases

The study involved an initial phase of establishing draft word lists for three semantic tasks matched for imageability, lexical frequency and word length, and collection of association norms from an online database. A second phase involved, novel normative data collection of semantic similarity ratings,[1] and visual similarity ratings (word-picture matching only). Within this phase additional items were substituted for any problematic items, and their ratings data collected afresh. This resulted in three word lists including semantic priming targets and primes, word-picture verification targets and distractors, and word-picture matching targets and semantic, phonological and unrelated distractors (see Table 1). We then sourced further data for these word lists from online sources, including values for concreteness, age of acquisition, and emotional valence.

Phase 1: construction of draft word lists

Selection of target stimuli

Three lists of 50 target nouns were developed, for which semantic distractors were identifiable. All the nouns were singular (with one exception *scissors*), and the majority were not homonyms or heteronyms (see Appendix A for exceptions to this). Where homonym or heteronym target words were included in the word-picture verification and word-picture matching tasks they were accompanied by the target picture, so the intended meanings were unambiguous. For words selected that were also verbs (e.g., map), in all cases except two the noun form had a higher lexical frequency than the verb form (see Appendix A for exceptions). Compound nouns were excluded, to prevent words within the compound inadvertently affecting semantic processing (e.g., chestnut). Words with negative emotional valence (e.g., coffin) were also excluded as individuals are reported to respond differently to these (e.g., Kensinger & Corkin, 2003). Items in the word-picture verification and word-picture matching tasks needed to be unambiguous in picture form. Targets in the word-picture matching task required a phonological neighbour which referred to a pictureable and semantically unrelated concept, such as target *mug* and distractor *slug*. The resulting three target word lists were initially matched for key psycholinguistic variables of imageability, lexical frequency and word length.[2] Imageability values were taken from the MRC Psycholinguistic database (Coltheart, 1981), and lexical frequency values were taken from the British National Corpus (The British National Corpus, version 3 (BNC XML Edition), 2007).

Table 1. Summary of novel semantic tasks and corresponding word lists.

Task	Target presentation	Matched semantic and associated pairs	Matched phonological distractors	Matched unrelated stimuli
Semantic priming task	Written	Written primes	N/A	Written
Word-picture verification task	Picture	Written distractors	N/A	N/A
Word-picture matching task	Written and picture	Picture distractors	Picture distractors	Picture distractors

Selection of semantic and associated partners

Semantic partners for the target words were selected; 32 targets were paired with a semantically similar word and 18 targets paired with an associated word. We derived the data for semantic association between targets and semantic partners from word association norms provided by the Edinburgh Association Thesaurus (Kiss et al., 1973). Association was measured in the direction of prime or distractor to target, on the basis that the prime or distractor appears first in semantic priming and word-picture verification, and this rule was generalised to word-picture matching.[3] The three lists of semantic partners met all the criteria outlined for the target words, with the exceptions detailed in Appendix A. The three semantic partner lists were controlled for psycholinguistic variables between lists, as per the target word lists criteria.

Selection of phonological and unrelated distractors for word-picture matching[4]

A list of 50 phonological distractor words was compiled for the word-picture matching test, with the criterion that the distractor either start or end with the same syllable as its target partner, or, for monosyllabic words, that at least 50% of the target phonemes were in the phonological partner word. The phonological distractors met the criteria outlined for targets and semantic partner words with exceptions shown in Appendix A.

Unrelated distractor words were selected and matched to the word-picture matching target words for imageability, lexical frequency and length. The unrelated words had no semantic or associative relationship to the targets and were checked to ensure minimum phonological or orthographic overlap with target partners. The unrelated distractors met the criteria used for all other lists outlined above for targets.

Draft word lists

The resulting word lists had the following properties. The three lists of target words were matched for imageability, lexical frequency and length, and the three lists of semantic and associative partners were also matched for these three key variables.[5]

Novel data

In phase two, novel data for the lists were collected from control participants. The data included semantic similarity ratings using a semantic similarity task for all word pairs used in the three tests (semantic priming: targets + semantic or associative primes; word-picture verification: targets + semantic or associative distractors; word-picture matching: targets + semantic or associative distractors; targets + phonological distractors; targets + unrelated distractors). Also, visual similarity ratings were collected using a visual similarity rating task for word-picture matching stimuli, which is the only test in which more than one image appears together.

Phase 2: normative data collection for test stimuli

Participants

Twenty adult participants were recruited to each of the normative data collection tasks of semantic similarity and visual similarity. Forty additional participants were recruited to complete extra data collection for a subset of word pairs that were replaced following the administration of the first rating tasks. The sample size was based on that used by two studies with similar experimental design (Cole-Virtue & Nickels, 2004a; Moss et al., 1995). Participants were recruited using opportunistic sampling methods via social networks. The inclusion and exclusion criteria for all participants were that individuals were aged 18 and over, monolingual literate native English speakers, with no history of speech, language or literacy impairment, no history of neurological disease, and sufficient visual acuity (aided or unaided) to enable accurate reading of written text. The gender and age data for the two groups of participants are shown in Table 2.

Ethical approval was granted by The University of Sheffield Department of Human Communication Sciences Research Ethics Committee. All participants were given a research information sheet and had the opportunity to ask questions. Participants read the information sheet then decided whether to continue to the task or not, with continuation with the task taken as informed consent to participate.

Materials

The semantic similarity task rating task addressed the semantic similarity of pairs of concepts expressed in written words and hence included 250 pairs of written single words. The visual similarity task addressed the visual similarity of pairs of concepts from the word-picture matching task, expressed in written words, and hence included 150 pairs of written single words. Rating tasks were completed either in paper form in person, or electronically via emailed materials.

Design

Participants were pseudo-randomly assigned to one similarity task only, with the provisos that twenty participants completed each task, roughly equal numbers of female and male participants completed each task, and that the age range across tasks was comparable.

In the semantic similarity task, the 250 word pairs were pseudo-randomly ordered to ensure that no more than three consecutive instances of the same category of relationship i.e., semantic, phonological or unrelated appeared. Targets from the word-picture matching task appeared in the task three times (i.e., with semantic, phonological and unrelated distractors), and each appearance of any one target word was separated by at least 10 intervening items. The 250 pairs were split into four blocks with four orders of presentation of the four blocks, order of presentation of blocks was then randomly assigned to participants. In the visual similarity task, the 150 words pairs (for the word-picture matching task only) were ordered as for the semantic similarity task. The 150 word

Table 2. Gender and age of participants.

Task	Number of participants	Female participants	Minimum age	Maximum age	Mean age
SST	20	12	25	48	32
VST	20	12	26	40	33
Additional SST data	20	11	22	49	32
Additional VST data	20	12	27	39	32

SST: semantic similarity task; VST: visual similarity task pairs were split into three blocks with four orders of presentation of each block and randomised as for the semantic similarity task.

Procedure

For both rating tasks, participants were provided with written instructions, a low and high similarity example pair, and a visual rating scale from 1 to 9. In the semantic similarity task, 1 was labelled not similar in meaning, 5 moderately similar in meaning, and 9 highly similar in meaning. In the visual similarity rating task, 1 was labelled not visually similar, 5 moderately visually similar, and 9 highly visually similar. In the semantic similarity task instructions, participants were informed: "Some words are very similar and related in **meaning**. For example, they may be from the same category such as furniture, animals or clothing." In the visual similarity task, participants were informed: "Objects can look visually similar or dissimilar, for example, they may be similar in appearance due to their size, shape and/or colour." Participants were instructed to consider each word pair in turn and provide a written number rating for each pair. For both tasks, participants were instructed to make their decisions independently, not spend too much time considering each word pair, and to complete the task in one sitting unless a break was needed. Participants generally completed the task within 20 minutes, however there was no time limit for completion. Participants had no further interaction with the researcher while completing the task, except where participants queried the meaning of homonyms.

The reliability of the semantic similarity scale was measured using Cronbach's alpha, within which values of over .7 are deemed to be acceptable. Analysis of all semantically similar pairs ($n = 96$) and of all associated pairs ($n = 54$) both gave Cronbach's $\alpha = .98$, demonstrating high reliability.

Additional data collection

Additional items replaced items with unsuitable ratings in the visual similarity task. Two items with the two highest visual similarity ratings were removed from the target-phonological and target-unrelated distractor categories.[6] Replacement stimuli were identified by matching a new item to the problematic item in terms of psycholinguistic variables. Where new pairs were introduced, new visual and semantic similarity ratings were collected by the additional group of 40 participants. The data presented in the results therefore represent the original stimuli, minus the problematic items, and with the addition of replacement items, ensuring that the final sets were matched for all variables under control.

Additional online data sourcing

Once the final word lists had been established, data concerning further variables were then collected from online resources hosted at the Center for Reading Research Website (http://crr.ugent.be). These included values for concreteness (Brysbaert et al., 2014), age of acquisition (Kuperman et al., 2012), and emotional valence (Warriner et al., 2013). These served to provide additional information about the sets and did not inform further changes to the lists.

Data analysis overview

The materials emerging from the processes described above consisted of the following: three lists of 50 target words, one for each semantic test; three lists of 50 semantic or associative partner words, one for each semantic test; and for the word-picture matching task, one list of phonologically related words and one list of unrelated words. In all three tests the target + semantic/associative partner pairs included 32 semantically similar pairs and 18 associated pairs.

The datasets included the following: targets in the three lists had a value for concreteness, imageability, age of acquisition, lexical frequency, word length, and emotional valence[7]; where available, distractor items across the three lists also had values for psycholinguistic variables. The target + semantic partner word pairs had values for their semantic similarity and association, as did the target + phonological distractor pairs and target + unrelated distractors in the word-picture matching task. The distractor words in the word-picture matching task had values for visual similarity between each target and each distractor. Where data were not attainable this is noted in the relevant table, with the number of cases provided.

Results

Tests of normality

Tests of normality (Kolmogorov Smirnov and Shapir Wilks tests), scrutiny of histograms, and measures of skew and kurtosis were used to investigate the distribution of each variable. Most of the data was non-normally distributed with no clear pattern allowing for transforming of datasets, therefore all statistical analysis is non-parametric. Kruskal-Wallis tests were used for comparisons involving three or more conditions, and Mann-Whitney U tests for two conditions. Where multiple post hoc comparisons were conducted the value of *p* was adjusted using Bonferroni corrections. Two-tailed significance levels are reported throughout unless stated otherwise.

Comparison of stimuli variables

Target words and semantic/associative partners: between-test comparison of variables

The values for psycholinguistic variables for the targets were compared between tests, and data are shown in Table 3(a). The corresponding data for the semantic and associative partners are shown in Table 3(b). No significant differences for any of the variables for either targets or semantic/associative partners were found.

Semantic similarity ratings and association values between targets and semantic/associated partners were also compared between the three tasks. Results are reported in Table 4. Analyses showed no significant differences in semantic similarity ratings or association values between tasks.

Word-picture matching: target and distractor within-test comparison

The word-picture matching target, semantic, phonological and unrelated distractors were compared in terms of psycholinguistic variables. Kruskall-Wallis tests were used to compare the sets with data shown in Table 5.

Table 3. (a) Target values for psycholinguistic variables: between-test comparisons (b) Semantic and associate partner values for psycholinguistic variables: between-test comparisons.

Psycholinguistic variable (n = 50 per set)	SP median	WPV median	WPM median	χ2	*df*	*p*
(a) Targets						
Concreteness	4.88 (0.18)	4.89 (0.14)	4.89 (0.28)	1.30	2	0.52
Imageability	595 (22.83)	597 (19.94)	597 (19.77)	0.55	2	0.76
Age of acquisition	5.40 (1.67)	5.11 (1.46)	5.06 (1.81)	0.90	2	0.64
Frequency	10.52 (30.36)	10.44 (25.51)	10.68 (25.74)	0.05	2	0.97
Letters	5 (1.49)	5 (1.54)	5 (1.48)	0.16	2	0.92
Phonemes	4 (1.41)	4 (1.41)	4 (1.51)	0.37	2	0.83
Syllables	1 (0.61)	1 (0.68)	1 (0.65)	0.33	2	0.86
Valence	5.88 (0.88) *49*	5.67 (0.95)	5.80 (1.05)	0.92	2	0.63
(b) Semantic and associate partners						
Concreteness	4.83 (0.33)	4.83 (0.15) *49*	4.89 (0.26) *46*	1.80	2	0.41
Imageability	595 (39.97) *37*	591 (37.94) *31*	591 (31.10) *32*	.99	2	0.61
Age of acquisition	6.00 (2.38) *49*	6.00 (2.16) *49*	5.30 (2.16) *46*	3.72	2	0.16
Frequency	6.38 (41.66)	5.58 (35.47)	9.95 (53.55)	0.53	2	0.77
Letters	5 (1.64)	5 (1.39)	5 (1.81)	0.22	2	0.89
Phonemes	4 (1.37)	4 (1.23)	4 (1.61)	0.53	2	0.77
Syllables	2 (0.58)	2 (0.63)	1.5 (0.73)	0.45	2	0.80
Valence	5.44 (0.96) *46*	5.52 (0.85) *48*	5.51 (1.10) *44*	0.93	2	0.63

Standard deviation in brackets. Number of values in a list is provided in italics where this is less than 50. SP: semantic priming; WPV: word-picture verification; WPM: word-picture matching.

Only concreteness showed a significant difference between word sets. Post hoc comparisons were conducted on this dataset using Mann-Whitney U tests, with Bonferroni correction of p to 0.00833. Only one comparison was significant: phonological

Table 4. Semantic similarity and association ratings between targets and partners: between-test comparisons.

Relationship rating (n = 50 per set)	SP median	WPV median	WPM median	χ2	*df*	*p*
Semantic similarity	5.65 (1.27)	5.60 (1.21)	5.38 (1.41)	1.96	2	0.38
Association	2.00 (8.73) *42*	2.00 (9.22) *43*	1.00 (9.20) *49*	.22	2	0.90

Standard deviation in brackets. Number of values in a list is provided in italics where this is less than 50. SP: semantic priming; WPV: word-picture verification; WPM: word-picture matching.

Table 5. Word-picture matching: comparison of psycholinguistic variables between targets and distractors.

Psycholinguistic variable	Targets median	Semantic/associative median	Phonological median	Unrelated median	χ2	*df*	*p*
Concreteness	4.89 (0.28)	4.89 (0.26) *46*	4.77 (0.27) *49*	4.87 (0.18)	13.01	3	<0.01
Imageability	597 (19.77)	591 (31.10) *32*	584 (40.74) *36*	592 (27.61)	5.49	3	0.14
Age of acquisition	5.06 (1.81)	5.30 (2.16) *46*	6.06 (1.55) *49*	5.48 (1.41)	7.11	3	0.07
Frequency	10.68 (25.74)	9.95 (53.55)	10.44 (63.95)	9.47 (22.55)	0.71	3	0.87
Letters	5 (1.48)	5 (1.81)	5 (1.44)	4.50 (1.72)	2.25	3	0.52
Phonemes	4 (1.51)	4 (1.61)	4 (1.29)	4 (1.33)	0.91	3	0.82
Syllables	1 (0.65)	1.5 (0.73)	1 (0.58)	1 (0.65)	1.99	3	0.57
Valence	5.80 (1.05)	5.51 (1.10) *44*	5.67 (1.01) *47*	5.72 (0.94)	0.88	3	0.83

Standard deviation in brackets. Number of values in the list is provided in italics where this is less than 50, where not provided all values are 50.

distractors had significantly lower concreteness values than targets ($U = 738.5$, $z = -3.410$, $p = .001$). All other comparisons were non-significant (targets vs semantic: $U = 970.5$, $z = -1.320$, $p = .188$ n.s.; targets vs unrelated: $U = 1065.5$, $z = -1.275$, $p = .204$ n.s.; semantic vs phonological: $U = 833.5$, $z = -2.189$, $p = .028$ n.s.; semantic vs unrelated: $U = 1132.0$, $z = -.132$, $p = .897$ n.s.; phonological vs unrelated: $U = 877.0$, $z = -2.439$, $p = .014$ n.s.).

Summary of word lists

Analysis of word list data showed that all critical variables were satisfactorily matched across the three tests, including psycholinguistic variables and semantic or associative relationship values. There was one potential problematic aspect, that phonological distractors had lower concreteness values than targets in word-picture matching. The phonological distractor set was retained however despite this, in order to maintain the phonological relationship between targets and distractors, as the latter was deemed of greater importance.

Semantically similarity and association measures: between-task comparison

Each target was paired with either a semantically similar pair ($n = 32$), or an associated pair ($n = 18$). The semantically similarity and association ratings of these pairs were compared between tests. The analyses demonstrate matching between tests (see Table 6), with no significant difference in semantic similarity or association for either the semantically similar or associated pairs.

Table 6. Semantic similarity and association values for semantically similar and associated pairs: between-test comparison.

	SP median	WPV median	WPM median	χ2	*df*	*p*
Semantically similar pairs						
Semantic similarity	6.20 (0.77)	6.18 (0.81)	5.98 (0.97)	1.33	2	0.51
Association	1.00 (2.74) *25*	1.00 (2.20) *27*	1.00 (3.26) *31*	0.82	2	0.67
Associated pairs						
Semantic similarity	3.95 (0.92)	4.00 (0.78)	3.73 (0.84)	3.19	2	0.20
Association	6.00 (12.29) *17*	14.00 (11.32) *16*	4.00 (13.63)	2.75	2	0.25

Rating scale for semantic similarity: 1 = not similar in meaning, 5 = moderately similar in meaning, 9 = highly similar in meaning. Standard deviations in brackets plus values of *n* in italics, where $n < 32$ (semantically similar pairs) or $n < 18$ (associated pairs) where association values were missing from the Edinburgh Association Thesaurus. SP: semantic priming; WPV: word-picture verification; WPM: word-picture matching.

Semantic similarity and association measures: within-test comparisons

The next analysis compared the pairs' semantic similarity ratings and association values within each test. Pairs had been selected as highly similar low-association pairs in the semantically similar group, and as highly associated low-similarity pairs in the associated group; the aim of these analyses was to verify this. The data are shown in Table 7, with Bonferroni adjustment applied and significance levels reported at 0.0167.

Table 7. Semantic similarity and association measures: within-test comparisons.

	SP Median	WPV Median	WPM Median
Semantic similarity ratings			
Semantically similar pairs (n = 32)	6.20 (0.77)	6.18 (0.81)	5.98 (0.97)
Associated pairs (n = 18)	3.95 (0.92)	4.00 (0.78)	3.73 (0.84)
U	29.5	34.5	28.0
z	−5.227	−5.125	−5.256
p	<0.001	<0.001	<0.001
Word association values			
Semantically similar pairs (n = 32)	1.00 (2.74) *25*	1.00 (2.20) *27*	1.00 (3.26) *31*
Associated pairs (n = 18)	6.00 (12.29) *17*	14.00 (11.32) *16*	4.00 (13.63)
U	106.5	55.0	186.5
z	−2.777	−4.111	−1.954
p	0.002	<0.001	0.026

Rating scale for semantic similarity: 1 = not similar in meaning, 5 = moderately similar in meaning, 9 = highly similar in meaning. Standard deviations in brackets plus values of *n* in italics, where $n < 32$ (semantically similar pairs) or $n < 18$ (associated pairs) where association values were missing from the Edinburgh Association Thesaurus. One tailed exact significance values of *p* reported as all comparisons predicted direction of difference. SP: semantic priming; WPV: word-picture verification; WPM: word-picture matching.

The data show that in all three tests the semantically similar pairs were rated significantly more similar than the associated pairs. Association values were significantly higher for the associated pairs than for the semantically similar pairs in the semantic priming and word-picture verification tests. Word-picture matching word association values were in the predicted direction but failed to reach significance.

Semantic similarity ratings of stimuli in word-picture matching

This analysis focused solely on word-picture matching, comparing the semantic similarity ratings of target-semantic/associated partners to the semantic similarity ratings between targets + phonologically related distractors, and targets + unrelated distractors that appear in the same array. Phonological and unrelated distractors had been selected to be semantically distant from the target presented simultaneously, and the analysis confirms this. Table 8 presents the participant median ratings of semantic similarity between target-semantic/associated, target-phonological and target-unrelated pairs in the word-picture matching test.

A Kruskall-Wallis test demonstrated that there was a significant difference in semantic similarity ratings between the distractor-target pairs in the word-picture matching task, $\chi^2(2) = 103.78$, $p < .000$. Pairwise comparisons were completed using Mann-Whitney U tests,

Table 8. Median semantic similarity ratings for target-distractor pairs in word-picture matching.

Semantic similarity ratings	Semantically similar/associated	Phonologically related	Unrelated
Number	50	50	50
Median	5.38	1.10	1.05
Standard deviation	0.82	0.24	0.22

Rating scale: 1 = not similar in meaning, 5 = moderately similar in meaning, 9 = highly similar in meaning.

with Bonferroni adjustment of p to 0.017 applied. The semantic/associative category was significantly more similar in meaning to targets than phonologically related items ($U = .000$, $z = -8.63$, $p < .001$), and unrelated items ($U = .000$, $z = -8.65$, $p < .001$), with the semantically similar/associated distractors rated as moderately similar in meaning to targets.

Visual similarity ratings of stimuli in word-picture matching

Participants' ratings from the visual similarity task are presented in Table 9. None of the pairs were rated as moderately or highly visually similar, which is important for the control of test visual stimuli.

Table 9. Comparison of visual similarity between stimuli in word-picture matching.

Visual similarity ratings	Target-semantic /associative distractors	Target-phonological distractors	Target-unrelated distractors
Number	50	50	50
Median	3.60	1.08	1.05
Standard deviation	2.18	0.53	0.35

Rating scale: 1 = not visually similar, 5 = moderately visually similar, 9 = highly visually similar.

A Kruskall-Wallis test comparing target-semantic, target-phonological and target-unrelated pair visual similarity ratings was significant, $\chi^2 = 59.21$, $p < .001$. Post hoc testing using Mann-Whitney U tests were carried out, with Bonferroni correction of *p* to 0.017 applied. There was no significant difference between phonological and unrelated pairs ($U = 1218$, $z = -.225$, $p = .824$), with both sets of distractors rated as not visually similar. However, semantic/associative pairs were rated as significantly more visually similar than phonological pairs ($U = 310.5$, $z = -6.506$, $p < .001$), and unrelated pairs ($U = 276$, $z = -6.749$, $p < .001$). This was due to higher visual similarity ratings for target-semantic distractor pairs (median = 4.93) compared to target-associated distractor pairs (median = 1.25) within the set of 50 pairs, and this difference was significant ($U = 17$, $z = 5.479$, $p < .001$).

The resulting word lists

The outcomes of the above processes resulted in three word lists to be used in semantic priming, word-picture verification and word-picture matching tests, with accompanying semantic and associated partner words to be used as primes (semantic priming) or distractors (word-picture verification, word-picture matching), and accompanying phonological and unrelated distractor words (word-picture matching). Each target word has a partner which is either semantically similar to or associated with the target, a variable which is not currently controlled for in semantic processing tests for people with aphasia and other neuro-cognitive conditions.

The sets were matched as follows: targets were matched across tasks for semantic, associative and psycholinguistic variables; semantic partners were matched across tasks for semantic, associative and psycholinguistic variables. The desired delineation between semantically similar and associated relationships was achieved in all three tests, with semantically similar word pairs higher in semantic similarity than associated word pairs, and associated word pairs higher in association than semantically similar word pairs. The word-picture matching task involves four images per trial and each distractor, phonological and unrelated item was checked for visual similarity to their target pair; only semantically similar partners were rated towards moderately visually similar to targets, whereas this was controlled for all other categories.

Discussion

The overall aim of this study was to provide novel materials with normative data for the purpose of enhancing the assessment of lexical semantics in people with aphasia and other neuro-cognitive conditions such as semantic dementia. To achieve that aim in phase one, three sets of target word stimuli were devised, which are matched on concreteness, imageability, frequency, word length, age of acquisition, and emotional valence. Three accompanying sets of semantic and associative partners were selected which are similarly matched across tasks. The relationship between targets and semantic/associative partners in terms of the above variables was found to be comparable across tasks. For the word-picture matching task, phonological distractors and unrelated words were matched to targets for the key psycholinguistic variables, except in the case of phonological distractors which were of lower concreteness than the target words. In phase two normative data regarding semantic similarity were collected, allowing the development of two subsets within each test; one subset where the target and semantic partner are semantically similar, and one subset where the target and semantic partner are associated, so that the potential impact of this variable can be separately investigated. Normative data regarding visual similarity were also collected for stimuli in the word-picture matching task and higher levels found in the target + semantically similar pairings (e.g., hamster-mouse), with all other pairs controlled for this variable. The tests represent an advance in methods from current explicit semantic tests, by controlling stimuli psycholinguistic variables, visual similarity, and semantic and associative stimuli relationships.

It is proposed that the use of these stimuli with populations with suspected lexical semantic impairment will enhance diagnosis. Individuals may make errors on word-picture verification and word-picture matching, but show positive effects of semantic priming, suggesting intact semantic knowledge. This was found in seminal studies in participants with fluent aphasia, who in explicit lexical comprehension tasks showed apparent lexical semantic impairment (e.g., Blumstein et al., 1982; Milberg & Blumstein, 1981; Milberg et al., 1987). These studies did not use matched stimuli across tests however so direct comparisons were not possible. Between-test matching in the current study allows the comparison of individuals' performance in tests which entail different levels of additional cognitive demands, and differing assessment of implicit (semantic priming) versus explicit processing (word-picture verification, word-picture matching). This type of future investigation will result in improved understanding of the nature of lexical semantic deficits, and the selection of more appropriate interventions, potentially leading to better rehabilitation outcomes for individuals.

Unlike other tests of lexical semantic processing, this study controlled for semantically similarity and association of word pairs in each task to allow investigations into the impact of type of relationship on processing. Individual differences are reported in the processing of similar versus associated (or "thematic") relations in neurologically unimpaired participants. In explicit similarity judgment tasks, some individuals show consistent matching for stimuli that share category relations whereas others match based on associative/thematic relation (Simmons & Estes, 2008). Eye-tracking methods have demonstrated that these differences between individuals remain across tasks with different requirements; in spoken word comprehension tasks where semantic stimuli act as distractors and are not explicitly considered in the task response, some individuals fixate more on thematic relations, others on taxonomic relations (Mirman & Graziano, 2012a). Additional research posits distinct neuroanatomical regions as the basis for differences between thematic versus categorical relation processing (Schwartz et al., 2011). Mirman and Graziano (2012a, 2012b) used eye tracking techniques to compare processing of stimuli presented with taxonomically related (semantically similar) versus thematic relationship (associated) distractors in a spoken word comprehension task. They found differences in performance between patients with posterior lesions and those with frontal lesions. The posterior lesion group performed better when pairs were semantically similar than when they were associated, which was not found in control or frontal lesion participants. The authors argue for distinct neuro-anatomical regions responsible for the two knowledge types. The word lists provided here allow investigation of the impact of semantic relationship type on semantic priming effect and word-picture verification and word-picture matching accuracy and response latencies, which will form novel contributions to the research in this area.

Visual similarity was also of concern due to its potential impact on processing ease, and in the current study visual similarity ratings were collected for the word pairs themselves. Similar to the confound in PALPA test stimuli, (Cole-Virtue & Nickels, 2004a), items rated as being highly semantically similar were also rated as being highly visually similar. This is not possible to resolve as highly semantically similar entities such as dog and cat typically share many semantic features (living, pet, has fur, has a tail), of which several are shared visual perceptual features. Future studies may benefit from using methods in which ratings of visual similarity are derived from participants rating test images rather than from the concepts represented by written word pairs. By providing values for the rated visual similarity, the matched word lists provide the possibility of probing the impact of this factor on processing in different populations, including those with visual-perceptual processing deficits.

One issue relating to the word sets remains unresolved. The phonological distractors and the targets in word-picture matching were not matched for concreteness. The criteria for identifying phonological overlap were used as the first condition for selection of these words and any differences in other variables had to be tolerated. The data are available to allow analysis of the impact of these factors on processing in any subsequent experiments involving these word lists.

The dataset has been further developed. Colour photographs and two novel colour images are incorporated for the word-picture verification and word-picture matching tasks that are uniform in size and presented with no or minimal contextual background. The semantic priming task has been developed into a continuous list paradigm with the requisite addition of filler words and non-words. Each task has been trialled in a computer software programme allowing automatic generation of accuracy and response latencies, which can then be used to examine test performance in control participants or people with suspected semantic impairment. The matched and controlled stimuli and accompanying data allow the comparison of performance across implicit and explicit tasks with differing cognitive demands, which has not been possible to date. The normative data also allow regression analysis for groups or individuals to identify predictor variables. As such the stimuli represent a novel contribution to neuropsychological testing in this domain and provide means for accurate diagnosis to support an appropriate choice of therapy for people with aphasia, ultimately aiming to improve rehabilitation outcomes. To date, the tests have been undertaken by 40 control participants and 20 people with aphasia. The semantic priming findings indicate that even in participants with aphasia who present with semantic impairment on explicit semantic judgment tasks, semantic activation is either fully retained, or operates typically then fails to decay (see Dyson et al., 2021 for full results). These findings bring to question the reliance on explicit, complex semantic tasks with the presence of competitors as diagnostic indicators in aphasia and other neurocognitive conditions.

Notes

1. The terms similar and associated are used throughout this report, with similar referring to two lexical concepts within the same semantic category and sharing many semantic features (e.g., rabbit and squirrel), and associated referring to two lexical concepts which co-occur frequently but are from different semantic categories and have no or few shared features (e.g., cow and grass).
2. At this stage the lists were draft only, so we do not report statistics. All lists underwent subsequent amendments following data collection from control participants and we report the statistical comparisons for the final word and image sets in the results section.
3. Note that the association rating between two words is different if measured in the opposite direction.
4. The semantic priming task also involved unrelated partners and to derive these the semantic partners were randomly reassigned to a target word and checked for semantic similarity or association.
5. Imageability data for the semantic and associative partners and phonological distractors were not all available. Partial data were therefore used to establish the sets. Concreteness values were later sourced for the final sets for all the words and matched across lists.
6. The target *arrow* and unrelated distractor *fork* were rated as 3.9, thus *fork* was replaced with flag. The target *kennel* and phonologically related *tunnel* were rated as 4.2, thus *tunnel* was replaced with *funnel*.
7. An emotional valence rating was not available for one target in the semantic priming word list.

Acknowledgments

The authors would like to thank the participants who carried out the stimuli ratings in this study. Special thanks also go to Dr Richard Body, who contributed to the word list selection, and to Jack Schofield, who contributed to some of the online data sourcing.

The authors have obtained permission to include Table 3(a,b) which appear as Tables 1 and 2 in the supplemental material file of Dyson, L., Morgan, J., Herbert, R. (2021). Semantic processing in aphasia: evidence from semantic priming and semantic interference. Language, Cognition and Neuroscience, *36*(4), 491-516, reprinted by permission of the publisher (Taylor & Francis Ltd, http://www.tandfonline.com).

Data availability statement

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

The assessment images are also available from the corresponding author, Lucy Dyson, upon reasonable request.

Disclosure of potential conflicts of interest

No potential conflict of interest was reported by the author(s).

Funding

This research was completed as part of a project funded by a Research Fellowship awarded to Lucy Dyson by the Stroke Association [grant number: TSAJRTF2011/01].

ORCID

Lucy Dyson http://orcid.org/0000-0002-6430-9731
Jane Morgan http://orcid.org/0000-0001-8024-028X
Ruth Herbert http://orcid.org/0000-0002-7139-1091

References

Bate, S., Kay, J., Code, C., Haslam, C., & Hallowell, B. (2010). Eighteen years on: What next for the PALPA? *International Journal of Speech-Language Pathology*, *12*(3), 190–202. https://doi.org/10.3109/17549500903548825

Baum, S. R. (1997). Phonological, semantic, and mediated priming in Aphasia. *Brain and Language*, *60*(3), 347–359. https://doi.org/10.1006/brln.1997.1829

Blumstein, S. E., Milberg, W., & Shrier, R. (1982). Semantic processing in aphasia: Evidence from an auditory lexical decision task. *Brain and Language*, *17*(2), 301–315. https://doi.org/10.1016/0093-934X(82)90023-2

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. https://doi.org/10.1016/S0028-3932(00)00034-8

Breese, E. L., & Hillis, A. E. (2004). Auditory comprehension: Is multiple choice really good enough? *Brain and Language*, *89*(1), 3–8. https://doi.org/10.1016/S0093-934X(03)00412-7

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Butterworth, B., Howard, D., & Mcloughlin, P. (1984). The semantic deficit in aphasia: The relationship between semantic errors in auditory comprehension and picture naming. *Neuropsychologia*, *22*(4), 409–426. https://doi.org/10.1016/0028-3932(84)90036-8

Chapman, C. A., Hasan, O., Schulz, P. E., & Martin, R. C. (2020). Evaluating the distinction between semantic knowledge and semantic access: Evidence from semantic dementia and comprehension-impaired stroke aphasia. *Psychonomic Bulletin & Review*, *27,* 607–639. https://doi.org/10.3758/s13423-019-01706-6

Cole-Virtue, J., & Nickels, L. (2004a). Spoken word to picture matching from PALPA: A critique and some new matched sets. *Aphasiology*, *18*(2), 77–102. https://doi.org/10.1080/02687030344000346

Cole-Virtue, J., & Nickels, L. (2004b). Why cabbage and not carrot?: An investigation of factors affecting performance on spoken word to picture matching. *Aphasiology*, *18*(2), 153–179. https://doi.org/10.1080/02687030344000517

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology*, *33A*(4), 497–505. https://doi.org/10.1080/14640748108400805

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*(9), 2593–2608. https://doi.org/10.1093/brain/awp146

Crutch, S. J., & Warrington, E. K. (2005). Gradients of semantic relatedness and their contrasting explanations in refractory access and storage semantic impairments. *Cognitive Neuropsychology*, *22*(7), 851–876. https://doi.org/10.1080/02643290442000374

Cutler, A. (1981). The reliability of speech error data. *Linguistics*, *19*(7-8), 561–582. http://dx.doi.org/10.1515/ling.1981.19.7-8.561

Dell, G. S., Martin, N., & Schwartz, M. F. (2007). A case-series test of the interactive two-step model of lexical access: Predicting word repetition from picture naming. *Journal of Memory and Language*, *56*(4), 490–520. https://doi.org/10.1016/j.jml.2006.05.007

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical Access in Aphasic and Nonaphasic Speakers. *Psychological Review*, *104*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Dyson, L., Morgan, J., & Herbert, R. (2021). Semantic processing in aphasia: Evidence from semantic priming and semantic interference. *Language, Cognition and Neuroscience*, *36*(4), 491-516. https://doi.org/10.1080/23273798.2020.1849749
Franklin, S. (1989). Dissociations in auditory word comprehension; evidence from nine fluent aphasic patients. *Aphasiology*, *3*(3), 189–207. https://doi.org/10.1080/02687038908248991
Franklin, S., Howard, D., & Patterson, K. (1995). Abstract word anomia. *Cognitive Neuropsychology*, *12*(5), 549–566. https://doi.org/10.1080/02643299508252007
Franklin, S., Turner, J., Lambon Ralph, M. A., Morris, J., & Bailey, P. J. (1996). A distinctive case of word meaning deafness? *Cognitive Neuropsychology*, *13*(8), 1139–1162. https://doi.org/10.1080/026432996381683
Harvey, D. Y., & Schnur, T. T. (2015). Distinct loci of lexical and semantic access deficits in aphasia: Evidence from voxel-based lesion-symptom mapping and diffusion tensor imaging. *Cortex*, *67*, 37–58. https://doi.org/10.1016/j.cortex.2015.03.004
Heaton, R., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtis, G. (1993). *Wisconsin card sorting test (WCST) manual revised and expanded*. Psychological Assessment Resources.
Heuer, S., & Hallowell, B. (2007). An evaluation of multiple-choice test images for comprehension assessment in aphasia. *Aphasiology*, *21*(9), 883–900. https://doi.org/10.1080/02687030600695194
Hillis, A. E., Rapp, B., Romani, C., & Caramazza, A. (1990). Selective impairment of semantics in lexical processing. *Cognitive Neuropsychology*, *7*(3), 191–243. https://doi.org/10.1080/02643299008253442
Hodges, J. R., Patterson, K., Oxbury, S., & Funnell, E. (1992). Semantic dementia: Progressive fluent aphasia with temporal lobe atrophy. *Brain*, *115*(6), 1783–1806. https://doi.org/10.1093/brain/115.6.1783
Howard, D., & Franklin, S. (1988). *Missing the meaning? A cognitive neuropsychological study of the processing of words by an aphasic patient*. MIT Press.
Howard, D., & Gatehouse, C. (2006). Distinguishing semantic and lexical word retrieval deficits in people with aphasia. *Aphasiology*, *20*(9), 921–950. https://doi.org/10.1080/02687030600782679
Howard, D., & Orchard-Lisle, V. (1984). On the origin of semantic errors in naming: Evidence from the case of a global aphasic. *Cognitive Neuropsychology*, *1*(2), 163–190. https://doi.org/10.1080/02643298408252021
Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A *micro*analytic review. *Psychonomic Bulletin & Review*, *10*(4), 785–813. https://doi.org/10.3758/BF03196544
Ivanova, M. V., & Hallowell, B. (2013). A tutorial on aphasia test development in any language: Key substantive and psychometric considerations. *Aphasiology*, *27*(8), 891–920. https://doi.org/10.1080/02687038.2013.805728
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*(5), 1065–1079. https://doi.org/10.1016/j.neuropsychologia.2006.09.009
Jefferies, E., & Lambon Ralph, M. (2006). Semantic impairment in stroke aphasia versus semantic dementia: A case-series comparison. *Brain*, *129*(8), 2132–2147. https://doi.org/10.1093/brain/awl153
Kay, J., Lesser, R., & Coltheart, M. (1992). *Psycholinguistic assessments of language processing in Aphasia*. Psychology Press.
Kay, J., Lesser, R., & Coltheart, M. (1996). Psycholinguistic assessments of language processing in aphasia (PALPA): An introduction. *Aphasiology*, *10*(2), 159–180. https://doi.org/10.1080/02687039608248403
Kensinger, E. A., & Corkin, S. (2003). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31*(8), 1169–1180. https://doi.org/10.3758/BF03195800
Kiss, G. R., Armstrong, C., Milroy, R., & Piper, J. (1973). An associative thesaurus of English and its computer analysis. *The Computer and Literary Studies*, 153–165.

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4
Lambon Ralph, M. A., Jefferies, E., Patterson, K., & Rogers, T. T. (2017). The neural and computational bases of semantic cognition. *Nature Reviews Neuroscience*, *18*(1), 42–55. https://doi.org/10.1038/nrn.2016.150
Lin, E. L., & Murphy, G. L. (2001). Thematic relations in adults' concepts. *Journal of Experimental Psychology: General*, *130*(1), 3–28. https://doi.org/10.1037/0096-3445.130.1.3
Lucas, M. (2000). Semantic priming without association: A meta-analytic review. *Psychonomic Bulletin & Review*, *7*(4), 618–630. https://doi.org/10.3758/BF03212999
Marinis, T. (2010). On-line sentence processing methods in typical and atypical populations. In S. Unsworth & E. Blom (Eds.), *Experimental methods in language acquisition research* (pp. 139–162). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.27.09mar
Martin, N., & Dell, G. S. (2019). Maintenance versus transmission deficits: The effect of delay on naming performance in aphasia. *Frontiers in Human Neuroscience*, *13*, 406. https://doi.org/10.3389/fnhum.2019.00406
McNamara, T. P., & Holbrook, J. B. (2003). Semantic memory and priming. In A. F. Healy & R. W. Proctor (Eds.), *Handbook of psychology: Experimental psychology* (Vol. 4, pp. 447–474). Wiley. https://doi.org/10.1002/0471264385.wei0416
McNamara, T. P. (2005). *Semantic priming: Perspectives from memory and word recognition*. Psychology Press. https://doi.org/10.4324/9780203338001
Meyer, D. E., & Schvaneveldt, R. W. (1971). Facilitation in recognizing pairs of words: Evidence of a dependence between retrieval operations. *Journal of Experimental Psychology*, *90*(2), 227–234. https://doi.org/10.1037/h0031564
Milberg, W., & Blumstein, S. (1981). Lexical decision and aphasia: Evidence for semantic processing. *Brain and Language*, *14*(2), 371–385. https://doi.org/10.1016/0093-934X(81)90086-9
Milberg, W., Blumstein, S. E., & Dworetzky, B. (1987). Processing of lexical ambiguities in aphasia. *Brain and Language*, *31*(1), 138–150. https://doi.org/10.1016/0093-934X(87)90065-4
Mirman, D., & Britt, A. E. (2014). What we talk about when we talk about access deficits. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1634), 20120388. http://dx.doi.org/10.1098/rstb.2012.0388
Mirman, D., & Graziano, K. M. (2012a). Individual differences in the strength of taxonomic versus thematic relations. *Journal of Experimental Psychology: General*, *141*(4), 601–609. https://doi.org/10.1037/a0026451
Mirman, D., & Graziano, K. M. (2012b). Damage to temporo-parietal cortex decreases incidental activation of thematic relations during spoken word comprehension. *Neuropsychologia*, *50*(8), 1990–1997. https://doi.org/10.1016/j.neuropsychologia.2012.04.024
Morris, J., & Franklin, S. (2012). Investigating the effect of a semantic therapy on comprehension in aphasia. *Aphasiology*, *26*(12), 1461–1480. https://doi.org/10.1080/02687038.2012.702885
Moss, H. E., Ostrin, R. K., Tyler, L. K., & Marslen-Wilson, W. D. (1995). Accessing different types of lexical semantic information: Evidence from priming. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21*(4), 863–883. https://doi.org/10.1037/0278-7393.21.4.863
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: Relationship between temporal lobe atrophy and semantic memory. *Annals of Neurology*, *47*(1), 36–45. https://doi.org/10.1002/1531-8249(200001)47:1%3C36::AID-ANA8%3E3.0.CO;2-L
Murray, L. L. (2002). Attention deficits in aphasia: Presence, nature, assessment, and treatment. *Seminars in Speech and Language*, *23*(2),107–116. https://doi:10.1055/s-2002-24987
Nestor, P. J., Fryer, T. D., & Hodges, J. R. (2006). Declarative memory impairments in Alzheimer's disease and semantic dementia. *Neuroimage*, *30*(3), 1010–1020. https://doi.org/10.1016/j.neuroimage.2005.10.008
Nickels, L., & Howard, D. (1995). Aphasic naming - What matters? *Neuropsychologia*, *33*(10), 1281–1303. https://doi.org/10.1016/0028-3932(95)00102-9

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 prefrontal and temporo-parietal cortices. *Journal of Cognitive Neuroscience*, *22*(7), 1597–1613. https://doi.org/10.1162/jocn.2009.21289

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. https://doi.org/10.1038/nrn2277

Rapp, B., & Caramazza, A. (2002). Selective difficulties with spoken nouns and written verbs: A single case study. *Journal of Neurolinguistics*, *15*(3), 373–402. https://doi.org/10.1016/S0911-6044(01)00040-9

Raven, J. C. (1956). *Coloured progressive matrices sets A, AB, B*. Oxford Psychologists Press Ltd.

Rose, T. A., Worrall, L. E., Hickson, L. M., & Hoffmann, T. C. (2012). Guiding principles for printed education materials: Design preferences of people with aphasia. *International Journal of Speech-Language Pathology*, *14*(1), 11–23. https://doi.org/10.3109/17549507.2011.631583

Rossion, B., & Pourtois, G. (2001). Revisiting Snodgrass and Vanderwart's object database: Color and texture improve object recognition. *Journal of Vision*, *1*(3), 413. https://doi.org/10.1167/1.3.413

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*(2), 217–236. https://doi.org/10.1167/1.3.413

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case series test of the two-step interactive model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, *54*(2), 228–264. https://doi.org/10.1016/j.jml.2005.10.001

Schwartz, M. F., Kimberg, D. Y., Walker, G. M., Brecher, A., Faseyitan, O. K., Dell, G. S., Mirman, D., & Coslett, H. B. (2011). Neuroanatomical dissociation for taxonomic and thematic knowledge in the human brain. *Proceedings of the National Academy of Sciences*, *108*(20), 8520–8524. https://doi.org/10.1073/pnas.1014935108

Shallice, T. (1987). Functional architecture of the language-processing system. In M. Coltheart., G. Sartori, & R. Job (Eds.), *The cognitive neuropsychology of language* (pp. 111–127). Lawrence Erlbaum Associates, Ltd.

Simmons, S., & Estes, Z. (2008). Individual differences in the perception of similarity and difference. *Cognition*, *108*(3), 781–795. https://doi.org/10.1016/j.cognition.2008.07.003

Tabor Connor, L., & Fucetola, R. P. (2011). Assessment of attention in people with aphasia: Challenges and recommendations. *Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders*, *21*(2), 55–63. https://doi.org/10.1044/nnsld21.2.55

The British National Corpus, version 3 (BNC XML Edition). (2007). *Distributed by Bodleian libraries, University of Oxford, on behalf of the BNC consortium*. http://www.natcorp.ox.ac.uk

Thompson, H. E., Robson, H., Lambon Ralph, M. A., & Jefferies, E. (2015). Varieties of semantic 'access' deficit in Wernicke's aphasia and semantic aphasia. *Brain*, *138*(12), 3776–3792. https://doi.org/10.1093/brain/awv281

Tulving, E., & Schacter, D. L. (1990). Priming and Human Memory Systems, *Science, 247*(4940), 301–306. https://doi.org/10.1126/science.2296719

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Warrington, E. K., & Cipolotti, L. (1996). Word comprehension: The distinction between refractory and storage impairments. *Brain*, *119*(2), 611–625. https://doi.org/10.1093/brain/119.2.611

Warrington, E. K., & McCarthy, R. (1983). Category specific access dysphasia. *Brain*, *106*(4), 859–878. https://doi.org/10.1093/brain/106.4.859

Warrington, E. K., & McCarthy, R. A. (1987). Categories of knowledge. *Brain*, *110*(5), 1273–1296. https://doi.org/10.1093/brain/110.5.1273

Warrington, E. K., & Shallice, T. (1979). Semantic access dyslexia. *Brain*, *102*(1), 43–63. https://doi.org/10.1093/brain/102.1.43

Yee, E. (2005). *The time course of lexical activation during spoken word recognition: evidence from unimpaired and aphasic individuals* [Unpublished doctoral dissertation]. Brown University.