Individual differences in verbal short-term memory and reading aloud: Semantic compensation for weak phonological processing across tasks

Nicola Savill[1,2], Piers Cornelissen[3] Junior Whiteley[2], Anna Woollams[4] Elizabeth Jefferies[2]

[1] York St John University

[2] University of York

[3] Northumbria University

[4] University of Manchester

Address for correspondence:

Nicola Savill

School of Psychological and Social Sciences, York St John University, York, YO31 7EX, UK

Keywords: semantic; phonological; verbal short-term memory; reading aloud; word repetition; individual differences

Abstract

According to contemporary accounts, linguistic behaviour reflects the interaction of distinct representations supporting word meaning and phonology. However, there is controversy about the extent to which this interaction occurs within task-specific systems, specialised for reading and short-term memory, as opposed to between components that support the full range of linguistic tasks. We examined whether individual differences in the efficiency of phonological processing would relate to the application of lexical-semantic knowledge to support verbal short-term memory, single word reading and repetition. In a sample of 83 participants we related nonword performance in each task (as a marker of phonological capacity in the absence of meaning) to the effects of word imageability (a lexical-semantic variable). We found stronger reliance on lexical-semantic knowledge in participants with weaker phonological processing. This relationship held across tasks, suggesting that lexical-semantic processing can compensate for phonological weakness which would otherwise give rise to poor performance. Our results are consistent with separable yet interacting primary systems for phonology and semantics, with lexical-semantic knowledge supporting pattern completion within the phonological system in a similar way across short-term memory and reading tasks.

Psycholinguistic and neuropsychological research indicates that language behaviours involve interactions between phonology and semantics (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Jefferies, Frankish, & Lambon Ralph, 2006; N. Martin & Saffran, 1997; Patterson & Lambon Ralph, 1999; Taylor, Duff, Woollams, Monaghan, & Ricketts, 2015; Tyler, Voice, & Moss, 2000), while neuroimaging studies reveal that these component processes are supported by distinct networks in the brain (e.g., Price, 2012; Saur et al., 2008). Largely parallel literatures have explored the predictions of this interactive architecture for verbal short-term memory (vSTM) and reading; however, if semantics and phonology correspond to primary systems for language processing, there should be similar relationships between tasks with quite different surface characteristics (i.e., linguistic behaviours driven by auditory and visual inputs) (Crisp & Lambon Ralph, 2006; Jefferies, Sage, & Lambon Ralph, 2008; Ueno, Saito, Saito, et al., 2011). In line with this view, the capacity to temporarily hold sequences of speech sounds in mind assists the comprehension and production of both spoken and written language (Daneman & Green, 1986; Jacquemot & Scott, 2006; McCarthy & Warrington, 1987; Pettigrew & Hillis, 2011; Vallar & Baddeley, 1987).

While influential models of vSTM have proposed a specialised phonological short-term store separate from (but linked to) long-term language systems (Atkinson & Shiffrin, 1968; Baddeley, 1986, 2000), alternative accounts, inspired by studies of patients with related language and STM impairment, propose that vSTM corresponds to temporary activation of long-term phonological/semantic knowledge (Acheson & MacDonald, 2009; N. Martin & Saffran, 1997; R. C. Martin, Lesch, & Bartha, 1999; Page, Madge, Cumming, & Norris, 2007). One language-based account (the 'semantic binding hypothesis'), originally proposed to account for immediate serial recall data from patients with semantic dementia, suggests that interactions between semantic and phonological representations are critical in stabilising phonological sequences

(Jefferies, Frankish, & Lambon Ralph, 2006; Jefferies, Grogan, Mapelli, & Isella, 2012; Knott, Patterson, & Hodges, 1997; Patterson, Graham, & Hodges, 1994). Semantic dementia patients show a progressive loss of semantic knowledge associated with atrophy and hypometabolism of the inferior anterior temporal lobes. In the context of relatively preserved phonological function (Jefferies, Jones, Bateman, & Lambon Ralph, 2005), they often make specific phonological errors during the short-term recall of words that are no longer fully understood (e.g., Jefferies, Hoffman, Jones, & Lambon Ralph, 2008; Jefferies, Jones, Bateman, & Lambon Ralph, 2004; Knott et al., 1997; Majerus, Norris, & Patterson, 2007; see also Peters, Majerus, De Baerdemaeker, Salmon, & Collette, 2009, for similar effects of semantic impairment in Alzheimer's Disease). Some converging support for the view that semantic information helps to stabilise the phonological trace has also been provided by studies of healthy participants (Hoffman, Jefferies, Ehsan, Jones, & Lambon Ralph, 2009; Jefferies, Frankish, & Lambon Ralph, 2006; Savill et al., 2018; Savill, Ellis, & Jefferies, 2017; Savill, Metcalfe, Ellis, & Jefferies, 2015).

Another common characteristic of semantic dementia is surface dyslexia – a marked deficit in reading exception words with uncommon spelling-to-sound correspondences (e.g., *yacht*) (Hodges, Patterson, Oxbury, & Funnell, 1992; Neary et al., 1998). Reading accuracy for exception words is strongly linked to the degree of semantic impairment in semantic dementia (Graham, Patterson, & Hodges, 2000; Jefferies, Lambon Ralph, Jones, Bateman, & Patterson, 2004; McKay, Castles, Davis, & Savage, 2007; Patterson et al., 2006; Woollams, Lambon Ralph, Plaut, & Patterson, 2007), consistent with the view that conceptual information makes an important contribution to the ability to generate the appropriate phonology for an orthographic input when the mapping between orthography and the phonological form is inconsistent with other similar words. Theoretical accounts of these data have parallels with the literature on verbal STM since, in both domains (single word reading; immediate serial recall), the proposal that semantic processing is a core component of the linguistic behaviour has been controversial (cf. dual route models of reading vs. interactive models of reading; Coltheart, Tree, & Saunders, 2010; Plaut, McClelland, Seidenberg, & Patterson, 1996; Woollams et al., 2007). 'Primary Systems' accounts of acquired language disorders, whereby difficulties across language tasks arise from disrupted interactions between underlying visual (e.g., orthographic), phonological and semantic representations (Patterson et al., 2006; Patterson & Lambon Ralph, 1999; Ueno, Saito, Rogers, & Lambon Ralph, 2011), resemble language-based explanations of short-term memory – i.e., both accounts envisage that semantic representations directly constrain phonological processing through their interaction.

The constraining influence of semantic representations may be more critical in situations where the required phonological patterns are unstable or weak, in both reading and immediate serial recall (i.e., for words with unusual spellings or in patients with phonological deficits). Patients with phonological deficits show stronger effects of semantic manipulations on single word reading performance, suggesting that they compensate for their phonological deficits using word meaning (e.g., Crisp & Lambon Ralph, 2006; Welbourne, Woollams, Crisp, & Lambon Ralph, 2011). A similar pattern has been found for repetition and immediate serial recall: phonologically-impaired patients show poor performance on nonwords and increased effects of lexical-semantic variables, such as imageability (Jefferies, Crisp, et al., 2006; Jefferies, Sage, et al., 2008; Katz & Goodglass, 1990; Verhaegen, Piertot, & Poncelet, 2013; Wilshire & Fisher, 2004). There is also considerable evidence for interactions between phonology and semantics in normal skilled reading: imageability effects on reading speed are more pronounced for words with inconsistent or exceptional spelling patterns, when the mapping between orthography and phonology is an unreliable guide to pronunciation (Cortese, Simpson, & Woolsey, 1997;

Hoffman, Lambon Ralph, & Woollams, 2015; Shibahara, Zorzi, Hill, Wydell, & Butterworth, 2003; Taylor et al., 2015; Woollams, 2005). Moreover, imageability effects for inconsistent/exception words are increased for individuals with relatively weak phonological skills (Strain & Herdman, 1999; Woollams, Lambon Ralph, Madrid, & Patterson, 2016). It is less straightforward to manipulate consistency in repetition and immediate serial recall (i.e., the ease of mapping from auditory inputs to articulation), although verbal short-term memory shows an influence of imageability (Allen & Hulme, 2006; Bourassa & Besner, 1994; Caza & Belleville, 1999; Jefferies, Frankish, & Ralph, 2006; Majerus & van der Linden, 2003; Miller & Roodenrys, 2009; Romani, McAlpine, & Martin, 2008; Tse & Altarriba, 2007; Walker & Hulme, 1999), as does the latency of single-item repetition (under manipulations of lexical competition or phonological typicality: Tyler et al., 2000; Ueno, Saito, Saito, et al., 2011). These effects are similar to findings from reading in that the influence of imageability tends to emerge under challenging conditions, such as under time pressure, high loads and/or phonological uncertainty.

Here we used an individual differences approach to further test the implications of this “primary systems” account. We reasoned that if phonology and lexical-semantic knowledge depend on separable yet interacting systems, individuals should vary on these components – some will have relative strengths in phonology, while others will have relative strengths in lexical-semantic aspects of language processing. Secondly, parallel effects of this variability should be seen across tasks with different superficial characteristics at both single and multiple word processing levels, such as verbal short-term memory (which involves auditory input) and word reading (which involves orthographic input). Thirdly, due to the interactive nature of these components, relative weakness in phonology (i.e., poorer performance on nonword performance), should lead to greater engagement of long-term lexical representations that might stabilise weak phonological patterns (giving rise to stronger effects of imageability). In this first study of normal variation in lexical-semantic effects in verbal short-term memory, we expected healthy participants with weaker phonological skills to show increased effects of imageability in all three tasks, mirroring patterns previously observed in reading (e.g., Hoffman et al., 2015; Strain & Herdman, 1999). Moreover, phonological performance in one task was expected to predict the effects of imageability in others.

We used imageability as a proxy for lexical-semantic knowledge, since the meanings of high imageability items are thought to be more readily accessible than the meanings of low imageability items (e.g., Plaut & Shallice, 1993). This variable is highly correlated with age-of-acquisition (AoA) and several studies have found that effects of imageability in reading are not independent of AoA (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Cortese & Khanna, 2007; Cortese & Schock, 2013; Monaghan & Ellis, 2002), although Shibahara et al., (2003) found an interaction between imageability and consistency in reading even when AoA was controlled. Modelling work suggests that AoA effects arise from the semantic pathway in reading (Lambon Ralph & Ehsan, 2006); however, debate about the nature of representations within this pathway (i.e., lexical vs. heteromodal concepts) is beyond the scope of this study. Consequently, we cannot interpret interactions between phonological skill and imageability as necessarily reflecting a role for heteromodal concepts, or for mental imagery. Instead, we assume that lexical-semantic knowledge is more readily accessed for imageable/early-acquired items, and then consider whether people with weaker phonological skills are more reliant on these representations in long-term memory. By using a robust manipulation that did not attempt to separate these highly-correlated constructs, we maximised the likelihood of detecting interactions with phonological skill.

Method

Participants

Eighty-three adult undergraduate students at the University of York, aged 18 to 27 years (M = 1.44 years; SD = 1.56), completed each of the tasks over two 1.5 hour testing sessions (over two days) in exchange for course credit or payment. All participants were native English speakers with normal hearing and normal or corrected-to-normal vision with no history of learning difficulties or language impairment. The study was approved by the Department of Psychology Ethics Committee at the University of York.

Experimental Stimuli

To maximise the likelihood of detecting imageability effects in our high-functioning, healthy adult participants, we used tasks and materials that should strongly tax phonological capacity. All words were low-to-medium frequency (according to SUBTLEX-UK Zipf frequencies: Van Heuven, Mandera, Keuleers, & Brysbaert, 2014; see Tables 1, 2 and 3 for average stimuli properties for the Immediate Serial Recall, Speeded Reading and Speeded Repetition tasks respectively; the full lists of items can be found in the Supplementary Materials). The length of lists for immediate serial recall were set at the limits of recall span or just beyond (6 and 7 words long, and 4 and 5 nonwords long; setting a high difficulty level for all trials).

(TABLE 1 ABOUT HERE)

We incorporated difficulty manipulations into the relatively simple tasks of single-item word reading and repetition to help reveal effects of lexical-semantic knowledge**.** Speeded reading contained a manipulation of spelling-sound consistency, since lexical-semantic factors in reading aloud are known to interact with consistency; inconsistent items show stronger benefits of imageability (Shibahara et al., 2003; Strain & Herdman, 1999; Strain, Patterson, & Seidenberg, 1995; Woollams, 2005; Woollams et al., 2016) – although as noted in the Introduction, these effects are not always found to be independent of AoA. In single-item repetition, we included a manipulation of phonotactic probability (driven by the biphone probability of the onset CV/VC), which influences both nonword and word repetition speeds (Vitevitch, 2003; Vitevitch & Luce, 1998, 2005). Phonological typicality modulates effects of imageability in this task (Ueno, Saito, Saito, et al., 2011) (example items are provided in Table 3). The facilitative influence of high phonotactic frequency is generally stronger for nonword than for word repetition, and it is increased for words when a sublexical processing strategy is employed (Vitevitch, Armbruster, & Chu, 2004; Vitevitch, Luce, Pisoni, & Auer, 1999; Vitevitch & Luce, 1999, 2004). Thus, in repetition tasks employing infrequent words, low phonotactic probability may increase effects of lexical-semantic knowledge; the phonological structure of the language does not readily support these items, similar to inconsistent spelling patterns in reading.

To provide a base level of stimulus control across our three experimental tasks (single-item reading, repetition and immediate serial recall) we created a set of 80 ‘core’ stimuli and mixed these with additional sets of stimuli that were specific to each task (these are shown in Table S1 in the Supplementary Materials). The core set consisted of 80 low frequency words with a CVC structure (where C = consonant, V = vowel) that were either high imageability (40 words with an imageability rating of >4 according to Cortese & Fugett, 2004) or low imageability (40 words with an imageability rating of <3.5 according to Cortese & Fugett, 2004). These high and low imageability words were selected so that (i) they could be further subdivided according to spelling-to-sound consistency for the single-item reading task (i.e., 20 with consistent spellings and 20 with inconsistent or irregular spellings within high and low imageability sets), (ii) they could otherwise be subdivided according to their phonotactic probability for the single-item repetition task 20 low and 20 high summed positional biphone probability words), and (iii) so that sets were matched for average word length (letters; all words were 3 phonemes in length), lexical frequency according to SUBTLEX-UK (van Heuven et al., 2014), word phonotactic frequency (Vitevitch & Luce, 2004), and average onset execution–acoustic intervals of the items' initial consonant sounds (to rule out any systematic articulatory influences on vocal RT in the single-item naming tasks; see Rastle, Croot, Harrington, & Coltheart, 2005). A set of 40 nonwords for use in each of the experimental tasks were created from the selected words by recombining the codas with bodies of other words and so that 20 of the new nonword spellings replicated the inconsistent letter combinations of the inconsistent words (e.g., 'worve', created from *worm* /wɜːm/ and *dove* /dʌv/, could potentially be pronounced /wɔːv/ or /wɜːv/).

(TABLE 2 ABOUT HERE)

An additional set of one hundred and sixty low frequency words were selected for the single-item reading task and another set of 160 low frequency words were selected for the single-item repetition task. Both of these sets consisted of 80 high and 80 low imageability words, of which half were monosyllabic and half were disyllabic (imageability ratings taken from Cortese & Fugett, 2004, and Schock, Cortese, & Khanna, 2012). In the reading set, the words were selected such that half had inconsistent or exception spellings (Glushko, 1979) while words in the repetition set were selected such that half had low phonotactic probability onsets. Forty monosyllabic and 40 disyllabic nonwords were created for use in the reading and repetition tasks in the same way as the core set of nonwords, such that spelling-to-sound consistency was manipulated for the reading task (i.e., 40 nonwords with an inconsistent digraph; 40 nonwords without ambiguous pronunciation) and onset biphone probabilities were manipulated for the repetition task (i.e., 40 nonwords with low onset biphone probabilities; 40 nonwords with common onset probabilities). Finally, CVC words were added to the 40 core high and low imageability words to create 20 high imageability and 20 low imageability lists for the ISR task (ten six and ten seven-item lists each) that avoided repetition of phonemes at each syllable position. Their phonemes were recombined within their respective lists to create nonwords for 20 lists (ten four and ten five-item lists).

(TABLE 3 ABOUT HERE)

Each stimulus was chosen to ensure that, within each task, high and low imageability sets remained matched for average lexical frequency (van Heuven et al., 2014), number of phonemes, number of letters, and summed biphone probability (Vitevitch & Luce, 2004), and so that the average interval between onset execution and audible sound for items in the single-item naming tasks were similar between conditions (Rastle et al., 2005).

With all stimuli combined, the imageability manipulations were matched across tasks (average word imageability ratings collapsed across condition: $F(2,737) = 0.15$, $p = .86$; reading $M = 4.30$, ISR $M = 4.33$, and repetition $M = 4.39$). Lexical frequency was also broadly similar between tasks (i.e., average frequencies all between a SUBTLEX Zipf value of 3 and 4) but words were of a significantly lower frequency on average in the reading task ($M = 3.24$) than in the immediate serial recall ($M = 3.49$) and repetition tasks ($M = 3.54$) ($F(2,737) = 23.63$, $p < .001$).

All spoken items were recorded by a native British female speaker in a soundproof booth and were edited in Praat to remove remaining noise. Monosyllabic items and disyllabic items were edited to 0.75 s and 0.8 s in length respectively.

Experimental Task Procedures

Immediate Serial Recall. In this task, participants attempted to repeat spoken word lists and nonword lists. They were asked to attempt to repeat all items immediately at the end of each list, in order of presentation, and to produce item attempts whenever possible, even if unsure. Trials were preceded by an instruction screen detailing the upcoming trial type (e.g., “6 words”, “4 nonwords”). The lists were presented auditorily, through a headset with integrated microphone, at a rate of one item per second. On screen, an exclamation mark appeared 200ms prior to the onset of the first spoken item and remained until the offset of the last item, when it was replaced by a question mark to cue the recall attempt. After verbally recalling each list, participants pressed a key to start the next trial. Responses were digitally recorded for later coding. Lists were presented in the same order for all participants: 20 six-item word lists (mixing 10 low imageability and 10 high imageability six-word lists), 10 four-nonword lists, followed by a rest break, 10 five-nonword lists and finally 20 seven-item lists (mixing 10 low imageability and 10 high imageability seven-word lists). This task took approximately 20 minutes on average to complete. Responses were phonemically transcribed offline, with each response coded (using a simplified version of the coding scheme used in Savill et al., 2015) as either target recalled correctly in position (CIP) or out of position (ORD); together, these responses captured items ‘correct in any position’. We used this as a metric of performance because this measure should more fully capture the lexical-semantic contribution to immediate serial recall, according to the literature (e.g., Poirier & Saint-Aubin, 1995). This is thought to be the case because when participants use lexical-semantic knowledge to maintain or regenerate the phonological trace, they frequently do not maintain the original serial position of the item. Non-target responses were coded as a phonologically related error (i.e. containing target phonemes in the correct syllable position; e.g., “cove” /kəʊv/ instead of ‘comb’), unrelated (i.e., it did not contain target phonemes in the correct syllable position), or as an omission error (in the case of a missing response).

Single-item Reading Task. Participants read aloud single words and nonwords as quickly and accurately as possible. Each item was centrally presented on screen in lowercase white Times New Roman font on a black background for 2700ms, during which time participants were expected to respond (responses were recorded until 3500ms after written word onset). Each item was preceded by a fixation cross shown for 300ms. Words and nonwords were presented in separate blocks of 60 items (high and low imageability words appeared in a fixed pseudo-random order within word blocks; we presented words and nonwords in separate blocks to maximise possible imageability effects on word RT) and instruction screens preceded each block informed the participant of the upcoming item types (i.e., ‘words’ or ‘nonwords’). Responses were digitally recorded as individual sound files in E-Prime. Spoken latencies and accuracy were manually coded offline using the CheckFiles program (Protopapas, 2007). Responses were marked incorrect if they did not correspond to the correct pronunciation, after allowing for accent variation and self-corrections (e.g., saying /bəʊ/ instead of /baʊ/ for *bough*; in the case of nonwords, responses were coded as inaccurate if they were not legal pronunciations, e.g., /məʊdʒ/ for *‘mogue’*). The task took 25 minutes to complete, on average.

Single-item Repetition Task. Participants repeated back auditorily presented single words and nonwords as quickly and accurately as possible. Each item was presented over headphones and participants could initiate their response at any point post audio onset. They had 3100 ms before the onset of the next trial (responses were recorded until 4000 ms post audio onset). Following the same structure as the reading task, words and nonwords were presented in blocks of 60 items (high and low imageability words in a fixed pseudo-random order within word blocks; again, we presented words and nonwords in separate blocks to maximise possible imageability effects on word RT) and instruction screens preceding each block informed the participant of the upcoming item types (i.e., ‘words’ or ‘nonwords’). Responses were digitally recorded as individual sound files in E-Prime. The task took, on average, 25 minutes to complete. Spoken latencies and accuracy were manually coded offline using the CheckFiles program (Protopapas, 2007). Responses were marked incorrect if they did not correspond to the correct pronunciation (acceptable pronunciations in the case of nonwords), after allowing for regional accent variation (errors were usually mumbled or truncated misarticulations).

Psychometric Measures. We ran computerised versions of measures of phonological skills and semantic knowledge: Our phonological measures were accuracy in the Spoonerism test from the York Adult Assessment Battery-Revised (Warmington, Stothard, & Snowling, 2013; this phonological manipulation task tests the spoonerised productions of auditorily-presented names, e.g., responding “Jichael Mackson” /dʒaɪ.kᵊl mæk.sᵊn/ to “Michael Jackson”) and the decoding efficiency subtest of the Test of Word Reading Efficiency (Torgesen, Wagner, & Rashotte, 1999; this tests the number of pseudowords correctly read aloud from a list within 45 seconds). Semantic knowledge was assessed using Warrington’s Graded Synonyms (Warrington, McKenna, & Orpwood, 1998; testing two alternative forced choice selections of appropriate synonyms for concrete and abstract words) and Warrington’s Picture Naming tests (Warrington, 1997; a picture naming test designed for adult populations).

Testing Protocol. Testing took place over two days: The order of the three experimental tasks (single-item reading, single-item repetition and immediate serial recall) was counterbalanced across participants: two of the three tasks were tested in the first session and the third was tested in the second session. The remaining measures were presented in the same order across the two sessions for all participants.

Results

Analysis strategy

(i) We start by presenting a linear mixed effects model of the experimental manipulations for each task separately, considering the effects of imageability or AoA on performance at the level of individual trials. This analysis suggested that imageability and AoA had equivalent effects; therefore, in subsequent stages, we used imageability as a proxy for lexical-semantic knowledge. (ii) Next, we considered individual differences in phonological skill within each task separately, using performance on nonwords as a marker of phonological ability. This analysis was performed on aggregated predicted data per subject and condition, derived from the first stage of the analysis. We found interactions between phonological ability and the effect of imageability in reading and ISR. (iii) The third analysis examined the relationship between phonological ability and imageability effects *across* as well as *within* tasks. We used a ratio score to take account of overall performance, since the magnitude of the imageability effect for a given numerical difference between high and low imageability items depends on the overall level of performance. (iv) In the final analysis, we used the same imageability ratio and examined its correlation with semantic and phonological skill, as assessed by standardised psychometric tests.

Models characterising the effect of imageability and age-of-acquisition (AoA) in each task

In the first step of the analysis, we characterised lexical-semantic effects in each task separately. We computed parallel models that used different, highly correlated markers of lexical-semantic knowledge – namely imageability and inverse AoA ratings (this correlation was $r = 0.68$, $p < .0001$ within our stimulus set; AoA ratings were taken from Kuperman, Stadthagen-Gonzalez, & Brysbaert (2012). In all models, we used continuous values of imageability and 1/AoA. We included all stimuli in these analyses (both the core set, repeated across tasks, and the stimuli uniquely presented in each task) (see below). The analysis was conducted using SAS 9.4.

For reading, we examined single-item reading times as the dependent measure, since there were relatively few errors. We used PROC MIXED to build a mixed linear model that tested the fixed effects of lexical-semantic knowledge (imageability or 1/AoA), consistency and their interaction, as well as stimulus exposure (i.e., whether it was the first, second or third time the stimulus had been presented to the participant, since the core subset was repeated across tasks). We also tested random effects of participant and item. Finally, we tested covariates relating to block (since the stimuli were presented in blocks separated by rest periods) and accuracy split by block and consistency (in order to control potential trade-offs between accuracy and reading times).

For repetition, we examined single-item response times as a marker of efficient phonological production. There were very few errors. We used PROC MIXED to build a mixed linear model that tested the fixed effects of lexical-semantic knowledge (imageability or 1/AoA), onset phonotactic probability and their interaction, as well as stimulus exposure (i.e., whether it was the first, second or third time the stimulus had been presented to the participant). We also tested random effects of participant and item. Finally, we tested covariates relating to block (since the stimuli were presented in blocks separated by rest periods) and accuracy split by block and consistency (in order to control potential trade-offs between accuracy and reading times).

For immediate serial recall, we used PROC GLIMMIX to build a generalised linear mixed model of single item recall, because this was an untimed task that yielded substantial numbers of errors. This approach allowed us to consider the effect of lexical-semantic knowledge item-by-item (incorrect ISR response categories could not be analysed in the same way and are summarised in the Supplementary Material). We assumed a binary distribution for the outcome and therefore used a logit link function. We tested the fixed effects of lexical-semantic knowledge (imageability or 1/AoA), serial position (i.e. the position of the item within the list), and list identity (i.e., each list to be repeated had a unique identifier, in order to account for variation in the difficulty of each list). We also tested random effects of participant and item.

Each model was optimized by ensuring that a) any fixed effect added to a model contributed a statistically significant reduction in -2 Log Likelihood, b) fixed effects were retained in a model only if their Type III test of fixed effects was significant at $p < .05$. The only exceptions to this were where one non-significant fixed effect comprised part of a significant two- or three-way interaction term, in which case it was retained. For the random effects in each model, we permitted individual variation at the intercept level for each participant/item, by including an unstructured variance-covariance matrix. The final models are shown in Table 4. For all models, we used the Satterthwaite method for computing the denominator degrees of freedom for the tests of fixed effects.

Figure 1 depicts the results of this analysis. The equivalence of imageability (left-hand column) and AoA (right-hand column; inverted such that high lexical-semantic knowledge is shown on the right-hand side of both graphs) is indicated by similarities in the size and direction of the estimated effects of imageability and 1/AoA. Note that item-level raw data used in analyses are accessible at goo.gl/ULYJgS.

INSERT FIGURE 1 ABOUT HERE

For the reading task, there were significant main effects of lexical-semantic knowledge (as indexed by imageability or AoA), spelling-sound consistency and stimulus exposure. There was also a significant consistency by lexical-semantic knowledge interaction.

In the repetition task, there were significant main effects of lexical-semantic knowledge (as indexed by imageability or AoA), stimulus exposure and block. There was no main effect of onset phonotactic probability in single-item repetition, and no interactions.

For the immediate serial recall task, there were significant effects of lexical-semantic knowledge (as indexed by imageability or AoA), serial position, and list identity.

All other effects for these models were non-significant.

INSERT TABLE 4 ABOUT HERE

Summary. All three language tasks showed significant effects of lexical-semantic knowledge, as indexed by either imageability or AoA ratings. Parallel analyses of imageability and AoA yielded identical outcomes, indicating that although the stimuli were selected on the basis of imageability ratings (see Methods), they were indexing lexical-semantic knowledge in general, and not specifically the ease of forming a mental image of each item. Lexical-semantic knowledge also interacted with spelling-sound consistency in reading aloud: words with inconsistent spelling-to-sound mappings showed larger effects of imageability and AoA, in line with previous findings (e.g., Cortese, Simpson, & Woolsey, 1997b; Monaghan & Ellis, 2002; Shibahara et al., 2003; Strain et al., 1995; Strain, Patterson, & Seidenberg, 2002; Woollams, 2005). Lexical-semantic knowledge did not interact with onset phonotactic probability.

Individual differences in the effects of imageability

The next analysis assessed the relationship between individuals' nonword performance and individual differences in the magnitude of the imageability effect in the same task. We examined aggregated RT data for reading and repetition, and aggregated accuracy data for immediate serial recall, using the models above to derive estimates of performance per trial per participant (and then averaging the trials in each condition). We used averaged data as opposed to individual trials, since our ultimate objective was to characterise the relationship between nonword performance and imageability effects *between* as well as within tasks -- and the tasks involved different types of performance metrics that could not readily be combined (RT and accuracy, modelled using linear and logistic models respectively).

Our experimental design included separate sets of 'high imageability' and 'low imageability' items, presented in different lists. The sets again included all items (not just the 'core' stimuli repeated across tasks). Both imageability and AoA variables were well-separated in these item sets, with non-overlapping confidence intervals (imageability: Low set *M* = 2.66, 99% CI [2.55– 2.78], High set *M* = 5.92, 99% CI [5.78 – 6.06]; AoA: Low set *M* = 11.17, 99% CI [10.67 – 11.67], High set *M* = 7.86, 99% CI [7.34 – 8.39]). We compared modelled estimates of performance on the high imageability and low imageability items using a mixed-effects model implemented using the PROC MIXED procedure in SAS (adopting a compound symmetry covariance structure). We included a covariate that related to nonword performance on the same task in each participant (average RT for nonword reading and repetition; average accuracy for nonword immediate serial recall), plus the interaction with imageability, allowing us to test whether the effect of imageability interacted with phonological skill as indexed by nonword performance.

INSERT FIGURE 2 ABOUT HERE

The modelled data for each task are depicted in Figure 2. In these plots, the red regression line shows performance on high imageability items, while the blue regression line shows performance on low imageability items. To aid interpretation, the black dashed line corresponds to the slope that would be required for the regression line for high imageability items to produce an imageability index (i.e. the ratio between high and low imageability scores) that was perfectly uncorrelated with nonword performance (see next section). It can be thought of as a ‘null’ line, and should not be confused with the presence or absence of a significant interaction between imageability and nonword performance. To illustrate, parallel regression lines for high and low imageability items will not produce a significant interaction term. However, parallel regression lines will produce an imageability index ratio that inevitably changes as a function of nonword performance.

This analysis for reading showed a main effect of imageability (*F*(1,80) = 11.15, *p* = .0013), a main effect of nonword performance on the same task (i.e., people who were slower to read nonwords were also generally slower to read words; *F*(1,80) = 253.98, *p* < .0001) and an interaction between these factors (*F*(1,80) = 63.82, *p* < .0001). Participants who were particularly slow to read nonwords were faster at reading high imageability words than would be predicted by the null line, consistent with our hypothesis (i.e. the high imageability regression line in red is less steep than the null line in Fig. 2a).

Repetition showed a main effect of imageability (*F*(1,80) = 36.20, *p* < .0001) and nonword performance (*F*(1,80) = 676.74, *p* < .0001) but no interaction between these factors (*F*(1,80) = 1.62, *p* = .21).

Immediate serial recall showed main effects of imageability (*F*(1,80) = 1225.12, *p* < .0001) and nonword performance (i.e., participants who were better at recalling nonwords also showed better performance for words; *F*(1,80) = 79.32, *p* < .0001). In addition, there was an interaction between these factors (*F*(1,80) = 52.91, *p* < .0001). Participants who were particularly good at recalling nonwords showed a weaker recall advantage for words that were high in imageability (relative to their performance on other words), so the red regression line lies below the null line in Fig. 2c. In other words, a larger proportion of the words recalled by participants with poor nonword immediate serial recall were high in imageability and consequently better supported by lexical-semantic knowledge.

Summary. There was a relationship between nonword performance and the effect of imageability in both reading and immediate serial recall, in line with our hypothesis. We did not observe this effect for repetition, perhaps because this task was relatively easy and showed the smallest effects of imageability overall.

Relationships with nonword performance across tasks

Having established a relationship between nonword performance and the magnitude of imageability effects for reading and immediate serial recall (although not repetition), we considered whether this relationship would hold across different language tasks. We used a ratio score to characterise the effects of imageability in each individual for reading and immediate serial recall: i.e. performance on low imageability items / performance on high imageability items. This ratio score produced positive numbers indexing imageability effects in reading, when RT was the dependent measure (i.e., individuals showing greater benefits for high imageability words showed greater deviation from zero in a positive direction). In contrast, imageability effects were indexed by negative numbers in immediate serial recall, when accuracy was the dependent measure (i.e., individuals showing greater benefits for high imageability words showed greater deviation from zero in a negative direction). The word repetition task was not included in this analysis, since it did not show an interaction between imageability and nonword performance above.

Nonword performance was correlated across reading, repetition and immediate serial recall, supporting the view that these tasks indexed phonological ability. Nonword repetition and reading RT showed a substantial correlation ($r = .458$, $p < .0001$). Nonword reading also correlated with nonword immediate serial recall performance ($r = -.293$, $p = .007$; this relationship was negative given these metrics involved RT and accuracy). There was no significant correlation between repetition and immediate serial recall for nonwords ($r = -.175$, $p = .113$).

Our main research question was whether the effect of imageability in reading and immediate serial recall would be greater for people with poor nonword performance across tasks.

These relationships are shown in Figure 3. As predicted, the imageability index for reading correlated with nonword performance; this relationship was significant in two of the three tasks. Participants who showed a larger reading advantage for high compared with low imageability items showed slower nonword reading (i.e., there was a significant positive correlation between the imageability index and nonword reading speed; $r = .644$, $p < .0001$; 95% confidence limits = .495 to .753). These participants were also slower to repeat nonwords ($r = .451$, $p = <.0001$; 95% confidence limits = .259 to .606). However, there was no relationship between the imageability index for reading and nonword immediate serial recall performance ($r = -.111$, $p = .31$). When we controlled for matrix reasoning performance, these correlations remained highly significant: a large effect of imageability in reading times was linked to long response times in nonword repetition ($r = .623$, $p < .0001$) and nonword reading ($r = .465$, $p < .0001$). There was still no correlation between the effect of imageability in reading and nonword immediate serial recall performance ($r = -.124$, $p = .27$).

INSERT FIGURE 3 ABOUT HERE

The imageability index for immediate serial recall correlated with nonword performance in all three tasks. Participants who showed a more substantial immediate serial recall advantage for high-imageability items (i.e. a stronger negatively-weighted imageability index) performed more poorly on nonword tasks, including lower accuracy on nonword immediate serial recall ($r = .706$, $p < .0001$; 95% confidence limits = .576 to .798), and longer response times on nonword repetition ($r = -.276$, $p = .01$; 95% confidence limits = -.463 to -.063) and nonword reading ($r = -.295$, $p = .007$; 95% confidence limits = -.478 to -.083). When we controlled for matrix reasoning performance, these correlations remained highly significant: a large effect of imageability in immediate serial recall was linked to low accuracy in nonword recall ($r = .706$, $p < .001$) and longer response times in nonword repetition (*r* = -.278, p = .01) and nonword reading (*r* = -.317, *p* = .004).

Summary. The effects of imageability in reading and immediate serial recall were related to nonword performance across tasks, as well as within tasks. Participants with better nonword performance showed a smaller effect of imageability.

Relationships with phonological and semantic test scores

In the final analysis, we assessed the relationship with performance on more standardised tests of phonology and semantics. We derived a phonological factor score from the average of TOWRE nonword reading and the spoonerism task z scores. These two phonological tasks were highly correlated, *r* = 0.50, *p* < .0001. We derived a semantic factor score from the average of the graded naming task and synonym judgement z scores, which were also highly correlated, *r* = 0.63, *p* < .0001. There was also a significant correlation between the semantic and phonological factor scores, reflecting the fact that people with good language skills tended to be good at both phonological and semantic tasks (*r* = 0.21, *p* = .045). For this reason, we regressed out semantic skills when considering the relationship between phonological performance and the imageability effect, and we regressed out phonological skills when examining the link between semantic performance and the imageability effect.

The imageability effect in word reading correlated with phonological performance (*r* = -.264, *p* = .02, 95% confidence limits = -.454 to -.048), while the relationship with semantic performance did not reach significance (*r* = -.116, *p* = .30). A larger effect of imageability in reading was linked to poor phonological skill, as predicted. The effect of imageability in immediate serial recall correlated with both phonological performance (*r* = .330, *p* = .003, 95% confidence limits = .120 to .509) *and* semantic performance (*r* = .330, *p* = .003, 95% confidence limits = .119 to .509). These analyses revealed that a larger effect of imageability was associated with relatively poor semantic as well as phonological skills. These partial correlations are shown in Figure 4.

Summary. Participants with poorer phonological skills showed larger imageability effects in reading and immediate serial recall, even when individual differences in semantic performance were taken into account, in line with our hypothesis. Semantic skills did not relate to the effects of imageability in reading, but they were associated with smaller imageability effects in immediate serial recall. This unexpected finding might have occurred because people with highly developed vocabularies were able to use lexical-semantic knowledge to support the immediate serial recall of low imageability as well as high imageability items; this would have reduced the influence of this variable in immediate serial recall.

FIGURE 4 ABOUT HERE

Discussion

Following three hours of testing with each of our 83 healthy and language-unimpaired adults, we found evidence that individual differences in primary systems thought to underpin language processing (i.e., phonology, lexical-semantic knowledge) interact to predict linguistic behaviour across multiple tasks with different surface characteristics (e.g., responses driven by auditory or visual inputs; involving single words or multiple items in a sequence).

(i) First, across reading, repetition and immediate serial recall of a sequence of items, we found consistent effects of imageability (a metric of lexical-semantic knowledge, which strongly correlates with age-of-acquisition). These effects were larger for words that had less consistent spellings (in reading), demonstrating that lexical-semantic knowledge contributes more to reading when the response is less constrained by orthographic input, in line with previous findings (Rodd, 2004; Shibahara et al., 2003; Strain & Herdman, 1999; Strain et al., 1995, 2002; Woollams, 2005; Woollams et al., 2016).

(ii) In addition, in reading and immediate serial recall, stronger effects of imageability were seen in participants with lower phonological skill, who performed more poorly on nonword versions of the tasks (since nonwords do not have corresponding meanings and are therefore a relatively pure measure of phonological processing capacity). Similar individual differences results have been reported for reading previously (Strain & Herdman, 1999; Woollams, Hoffman, Roberts, Lambon Ralph, & Patterson, 2014; Woollams et al., 2016). However, studies of semantic contributions to verbal short-term memory have focused on effects arising at the group level (single word repetition: Tyler et al., 2000; Ueno, Saito, Saito, et al., 2011; immediate serial recall: Acheson, Postle, & MacDonald, 2010; Jefferies, Frankish, & Lambon Ralph, 2006; Majerus & van der Linden, 2003; Savill et al., 2018, 2017, 2015) and have not considered how these influences may vary across individuals as a function of phonological capacity, as we test here.

(iii) Finally, a particular strength of this research is that multiple language tasks were conducted in the same sample, allowing us to test the hypothesis arising from the primary systems account that underlying phonological skill assessed using nonword capacity in one task would predict the reliance on lexical-semantic knowledge in other language tasks. This prediction was borne out in relationships between reading aloud and verbal short-term memory performance (but not at an individual level in single item repetition, possibly due to weak sensitivity to

> variation in performance in this task, which provided participants with an immediately available model of the correct response and was therefore easy to perform). We conclude that in phonologically-challenging experimental conditions (e.g., span/supra-span length lists in immediate serial recall, words with inconsistent spellings in reading), word meaning supports performance in a comparable way across tasks, particularly in individuals with weaker phonological processing skills.

These findings are highly consistent with theories derived largely on the basis of neuropsychological data. Patients who have phonological deficits following stroke aphasia show increased effects of semantic variables, such as imageability, in language tasks including reading and repetition (Crisp, Howard, & Lambon Ralph, 2011; Crisp & Lambon Ralph, 2006; Hanley & Kay, 1997; Jefferies, Crisp, et al., 2006). For example, in a comparative study of semantic dementia and phonological dyslexia, Jefferies, Crisp, & Lambon Ralph (2006) reported that the size of the imageability effect in repetition correlated with the degree of phonological deficits and not semantic impairment. Similar findings have been found for reading (Crisp et al., 2011; Crisp & Lambon Ralph, 2006; Roberts, Lambon Ralph, & Woollams, 2010) and immediate serial recall (Wilshire, Keall, & O’Donnell, 2010). Our large-scale individual differences study provides important corroboratory evidence that an interaction between lexical-semantic and phonological representations underpins these language tasks, given the small sample size of some of these patient studies and the possibility of additional, undocumented, deficits. Neuropsychological studies also show that, just as phonological weakness increases the sensitivity to semantic effects, semantic impairment can elicit phonological errors on tasks such as reading and immediate serial recall (Graham et al., 2000; Hoffman et al., 2009; Jefferies, Frankish, & Noble, 2009; Jefferies, Hoffman, et al., 2008; Knott et al., 1997; McKay et al., 2007;

Patterson et al., 2006, 1994; Woollams, 2015; Woollams et al., 2007). These errors are thought to reflect a failure to constrain phonological production in the absence of strong constraints from the semantic system (Jefferies et al., 2005; Patterson et al., 1994). They occur under similar circumstances to those shown to be important for the interaction between phonology and semantics in the current study (e.g., the repetition of multiple items in immediate serial recall; reading aloud words with inconsistent spellings).

More broadly, these findings support accounts of language processing that anticipate a trade-off between semantic and phonological processes – both of which are thought to contribute to word processing within an interactive framework (e.g., Foygel & Dell, 2000; D.C. Plaut & Kello, 1999). These accounts make the prediction that a similar level of performance on word tasks may be achieved by participants who have relative strengths in either phonological or semantic processing – while these groups might be *quantitatively* similar in their overall performance, they should be *qualitatively* distinct. For example, participants relying more on phonological processes will be less influenced by word imageability and perform better on items that deviate from the long-term phonological structure within the language, in line with our findings. Our results uniquely show that assessments of phonological capacity in one task relate to effects of lexical-semantic knowledge in another, consistent with the view that these cross-participant differences emerge from differences in the recruitment of primary phonological and semantic systems.

There might be alternative ways of explaining our results. First, in order to maximise the interaction between lexical-semantic knowledge and phonological skill, we did not separate lexical and conceptual contributions to the imageability effect. Imageability relates to the availability of concrete semantic features during word processing (Plaut & Shallice, 1993), and correlates highly with AoA (Cortese & Schock, 2013; Gilhooly & Logie, 1980; Monaghan &

Ellis, 2002). Consequently, it is unclear if people with weaker phonological skill draw more on heteromodal aspects of conceptual knowledge or more specifically on lexical-semantic knowledge. Future work could attempt to separate these alternatives experimentally (i.e., via non-verbal priming of conceptual representations). Secondly, although we controlled for IQ in the analysis, we remain open to the possibility that there could be basic facets of cognition that underpin individual differences in lexical-semantic and phonological domains. For example, phonological tasks require the maintenance and manipulation of externally-presented information, with potentially a smaller role for retrieval from long-term memory in general terms, while semantic tasks are sensitive to the availability and accessibility of information in long-term memory, with potentially a smaller role for on-line processing. Thirdly, “division of labour” explanations are hard to distinguish from “diminishing returns”. By the first view, individuals and specific items differentially rely on semantics and phonology, reflecting the strengths and weaknesses of these processes. Alternatively, the second view predicts that if phonology is very strongly engaged, there is little potential for semantic variables to impact on performance (cf. Plaut et al., 1996), even if semantic processes are not inherently weak. We cannot strongly separate these accounts since we relate nonword processing to effects of imageability in language tasks that also draw strongly on phonology.

Considering the common behavioural relationships between phonological and semantic function in reading and verbal short-term memory, we might expect commonalities in the neural bases of semantic reliance in these tasks. Thus, future research might examine whether functional variation corresponding to the influence of semantics in unimpaired reading aloud would extend to other language contexts such as verbal short-term memory/repetition: for example, individual differences in structural connectivity between areas that process semantic and phonological information (cf. Graves et al., 2014) and/or differential activation in semantic and phonological regions (cf. Hoffman et al., 2015). The results also have practical implications. Since even a relatively homogeneous sample of participants (i.e., high-achieving and young undergraduates) show variability in the underlying systems that are employed to support language tasks, we anticipate that factors such as age, individual differences in the strength of intrinsic connectivity in different brain networks, psychoactive drugs and patterns of development will modulate the cognitive architecture that underpins language, even when no differences are seen in performance (Davies, Arnell, Birchenough, Grimmond, & Houlson, 2017; Dilkina, McClelland, & Plaut, 2008).

In conclusion, this work was inspired by two independent areas of research: 1) recent theoretical accounts of verbal short-term memory which hold that phonological maintenance is influenced by the same long-term linguistic and semantic knowledge that shapes everyday spoken language production and comprehension, contrary to influential views of verbal STM that propose a dedicated short-term storage system (‘the phonological loop’) and 2) a growing literature on individual differences in reading aloud, which has shown that sensitivity to variables such as imageability reflect the ease of phonological processing. We have demonstrated the connection between these two bodies of literature by showing that there are parallel individual differences in reading and verbal STM, which reflect the increased reliance on lexical-semantic representations in participants who have weaker phonological skills and in circumstances when the phonological pattern required by the task is less consistent with long-term linguistic experience.

Funding

This work was funded by a European Research Council grant (283530 – SEMBIND).

References

Acheson, D. J., & MacDonald, M. C. (2009). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin*, *135*(1), 50–68. https://doi.org/10.1037/a0014411.

Acheson, D. J., Postle, B. R., & MacDonald, M. C. (2010). The interaction of concreteness and phonological similarity in verbal working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(1), 17–36. https://doi.org/10.1037/a0017679.The

Allen, R. J., & Hulme, C. (2006). Speech and language processing mechanisms in verbal serial recall. *Journal of Memory and Language*, *55*(1), 64–88. https://doi.org/10.1016/j.jml.2006.02.002

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In *The psychology of learning and motivation (Vol 2)*. London: Academic Press.

Baddeley, A. D. (1986). *Working Memory*. Oxford, UK: Clarendon Press.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283–316. https://doi.org/10.1037/0096-3445.133.2.283

Bourassa, D. C., & Besner, D. (1994). Beyond the articulatory loop: A semantic contribution to serial order recall of subspan lists. *Psychonomic Bulletin & Review*, *1*(1), 122–125. https://doi.org/10.3758/BF03200768

Caza, N., & Belleville, S. (1999). Semantic contribution to immediate serial recall using an unlimited set of items: evidence for a multi-level capacity view of short-term memory. *International Journal of Psychology*, *34*(5–6), 334–338. https://doi.org/10.1080/002075999399657

Coltheart, M., Tree, J. J., & Saunders, S. J. (2010). Computational modeling of reading in semantic dementia: comment on Woollams, Lambon Ralph, Plaut, and Patterson (2007). *Psychological Review*, *117*(1), 256–72. https://doi.org/10.1037/a0015948

Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 384–387. https://doi.org/10.3758/BF03195585

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *Quarterly Journal of Experimental Psychology*, *60*(8), 1072–1082. https://doi.org/10.1080/17470210701315467

Cortese, M. J., & Schock, J. (2013). Imageability and age of acquisition effects in disyllabic word recognition. *Quarterly Journal of Experimental Psychology*, *66*(5), 946–972. https://doi.org/10.1080/17470218.2012.722660

Cortese, M. J., Simpson, G. B., & Woolsey, S. (1997). Effects of association and imageability on phonological mapping. *Psychonomic Bulletin & Review*, *4*(2), 226–31. https://doi.org/10.3758/BF03209397

Crisp, J., Howard, D., & Lambon Ralph, M. A. (2011). More evidence for a continuum between phonological and deep dyslexia: Novel data from three measures of direct orthography-to-phonology translation. *Aphasiology*, *25*(5), 615–641. https://doi.org/10.1080/02687038.2010.541470

Crisp, J., & Lambon Ralph, M. A. (2006). Unlocking the nature of the phonological-deep dyslexia continuum: the keys to reading aloud are in phonology and semantics. *Journal of Cognitive Neuroscience, 18*(3), 348–62. https://doi.org/10.1162/089892906775990543

Daneman, M., & Green, I. (1986). Individual Differences in Comprehending in Context and Producing Words. *Journal of Memory and Language*, *25*, 1–18. https://doi.org/10.1016/0749-596X(86)90018-5

Davies, R., Arnell, R., Birchenough, J. M. H., Grimmond, D., & Houlson, S. (2017). Reading through the lifespan: Individual differences in psycholinguistic effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*(8), 1298–1338. https://doi.org/10.1037/xlm0000366

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

Dilkina, K., McClelland, J. L., & Plaut, D. C. (2008). A single-system account of semantic and lexical deficits in five semantic dementia patients. *Cognitive Neuropsychology*, *25*(2), 136–64. https://doi.org/10.1080/02643290701723948

Foygel, D., & Dell, G. S. (2000). Models of Impaired Lexical Access in Speech Production. *Journal of Memory and Language*, *43*(2), 182–216. https://doi.org/10.1006/jmla.2000.2716

Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation, 12*(4), 395–427. https://doi.org/10.3758/BF03201693

Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance, 5*(4), 674–691. https://doi.org/10.1037/0096-1523.5.4.674

Graham, N. L., Patterson, K. E., & Hodges, J. R. (2000). The impact of semantic memory impairment on spelling: evidence from semantic dementia. *Neuropsychologia*, *38*, 143–163. https://doi.org/10.1016/S0028-3932(99)00060-3

Graves, W. W., Binder, J. R., Desai, R. H., Humphries, C., Stengel, B. C., & Seidenberg, M. S. (2014). Anatomy is strategy: Skilled reading differences associated with structural connectivity differences in the reading network. *Brain and Language*, *133C*, 1–13. https://doi.org/10.1016/j.bandl.2014.03.005

Hanley, J. R., & Kay, J. (1997). An Effect of Imageability on the Production of Phonological Errors in Auditory Repetition. *Cognitive Neuropsychology*, *14*(8), 1065–1084. https://doi.org/10.1080/026432997381277

Hodges, J. R., Patterson, K. E., 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

Hoffman, P., Jefferies, E., Ehsan, S., Jones, R. W., & Lambon Ralph, M. A. (2009). Semantic memory is key to binding phonology: converging evidence from immediate serial recall in semantic dementia and healthy participants. *Neuropsychologia*, *47*(3), 747–760. https://doi.org/10.1016/j.neuropsychologia.2008.12.001

Hoffman, P., Lambon Ralph, M. A., & Woollams, A. M. (2015). Triangulation of the neurocomputational architecture underpinning reading aloud. *Proceedings of the National Academy of Sciences*, *112*(28), E3719–E3728. https://doi.org/10.1073/pnas.1502032112

Jacquemot, C., & Scott, S. K. (2006). What is the relationship between phonological short-term memory and speech processing? *Trends in Cognitive Sciences*, *10*(11), 480–486. https://doi.org/10.1016/j.tics.2006.09.002

Jefferies, E., Crisp, J., & Lambon Ralph, M. A. (2006). The impact of phonological or semantic impairment on delayed auditory repetition: Evidence from stroke aphasia and semantic dementia. *Aphasiology*, *20*(9), 963–992. https://doi.org/10.1080/02687030600739398

Jefferies, E., Frankish, C., & Noble, K. (2009). Lexical coherence in short-term memory: strategic reconstruction or “semantic glue”? *Quarterly Journal of Experimental Psychology*, *62*(10), 1967–82. https://doi.org/10.1080/17470210802697672

Jefferies, E., Frankish, C. R., & Lambon Ralph, M. A. (2006). Lexical and semantic binding in verbal short-term memory. *Journal of Memory and Language*, *54*(1), 81–98. https://doi.org/10.1016/j.jml.2005.08.001

Jefferies, E., Frankish, C., & Ralph, M. A. L. (2006). Lexical and semantic influences on item and order memory in immediate serial recognition: Evidence from a novel task. *Quarterly Journal of Experimental Psychology*, *59*(5), 949–964. https://doi.org/10.1080/02724980543000141

Jefferies, E., Grogan, J., Mapelli, C., & Isella, V. (2012). Paced reading in semantic dementia: word knowledge contributes to phoneme binding in rapid speech production. *Neuropsychologia*, *50*(5), 723–32. https://doi.org/10.1016/j.neuropsychologia.2012.01.006

Jefferies, E., Hoffman, P., Jones, R., & Lambon Ralph, M. A. (2008). The impact of semantic impairment on verbal short-term memory in stroke aphasia and semantic dementia: A comparative study. *Journal of Memory and Language*, *58*(1), 66–87. https://doi.org/10.1016/j.jml.2007.06.004

Jefferies, E., Jones, R., Bateman, D., & Lambon Ralph, M. A. (2004). When does word meaning affect immediate serial recall in semantic dementia? *Cognitive, Affective & Behavioral Neuroscience*, *4*(1), 20–42. https://doi.org/10.3758/CABN.4.1.20

Jefferies, E., Jones, R. W., Bateman, D., & Lambon Ralph, M. A. (2005). A semantic contribution to nonword recall? Evidence for intact phonological processes in semantic dementia. *Cognitive Neuropsychology*, *22*(2), 183–212. https://doi.org/10.1080/02643290442000068

Jefferies, E., Lambon Ralph, M. A., Jones, R., Bateman, D., & Patterson, K. (2004). Surface dyslexia in semantic dementia: A comparison of the influence of consistency and regularity. *Neurocase*, *10*(4), 290–299. https://doi.org/10.1080/13554790490507623

Jefferies, E., Sage, K., & Lambon Ralph, M. A. (2008). Do deep dyslexia, dysphasia and dysgraphia share a common phonological impairment? *Neuropsychologia*, *45*(7), 1553–1570. https://doi.org/10.1016/j.neuropsychologia.2006.12.002

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: is it phonological and is it a store? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(3), 656–74. https://doi.org/10.1037/0278-7393.30.3.656

Katz, R. B., & Goodglass, H. (1990). Deep dysphasia: Analysis of a rare form of repetition disorder. *Brain and Language*, *39*(1), 153–185. https://doi.org/10.1016/0093-934X(90)90009-6

Knott, R., Patterson, K. E., & Hodges, J. R. (1997). Lexical and semantic binding effects in short-term memory: Evidence from semantic dementia. *Cognitive Neuropsychology*, *14*(8), 1165–1216. https://doi.org/10.1080/026432997381303

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., & Ehsan, S. (2006). Age of acquisition effects depend on the mapping between representations and the frequency of occurrence: Empirical and computational evidence. *Visual Cognition*, *13*(7–8), 928–948. https://doi.org/10.1080/13506280544000110

Majerus, S., Norris, D. G., & Patterson, K. E. (2007). What does a patient with semantic dementia remember in verbal short-term memory? Order and sound but not words. *Cognitive Neuropsychology*, *24*(2), 131–151. https://doi.org/10.1080/02643290600989376

Majerus, S., & van der Linden, M. (2003). Long-term memory effects on verbal short-term memory: A replication study. *British Journal of Developmental Psychology*, *21*, 303–310. https://doi.org/10.1348/026151003765264101

Martin, N., & Saffran, E. M. (1997). Language and auditory-verbal short-term memory impairments: Evidence for common underlying processes. *Cognitive Neuropsychology*, *14*(5), 641–682. https://doi.org/10.1080/026432997381402

Martin, R. C., Lesch, M. F., & Bartha, M. C. (1999). Independence of input and output phonology in word processing and short-term memory. *Journal of Memory and Language*, *41*(1), 3–29. https://doi.org/10.1006/jmla.1999.2637

McCarthy, R. A., & Warrington, E. K. (1987). Understanding: a Function of Short-Term Memory? *Brain*, *110*(6), 1565–1578. https://doi.org/10.1093/brain/110.6.1565

McKay, A., Castles, A., Davis, C., & Savage, G. (2007). The impact of progressive semantic loss on reading aloud. *Cognitive Neuropsychology*, *24*(2), 162–186. https://doi.org/10.1080/02643290601025576

Miller, L. M., & Roodenrys, S. (2009). The interaction of word frequency and concreteness in immediate serial recall. *Memory & Cognition*, *37*(6), 850–65. https://doi.org/10.3758/MC.37.6.850

Monaghan, J., & Ellis, A. W. (2002). What exactly interacts with spelling--sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 183–206. https://doi.org/10.1037//0278-7393.28.1.183

Neary, D., Snowden, J. S., Gustafson, L., Passant, U., Stuss, D., Black, S. E., … Benson, D. F. (1998). Frontotemporal lobar degeneration: A consensus on clinical diagnostic criteria. *Neurology*, *51*(6), 1546–1554. https://doi.org/10.1212/WNL.51.6.1546

Page, M. P. A., Madge, A., Cumming, N., & Norris, D. G. (2007). Speech errors and the phonological similarity effect in short-term memory: Evidence suggesting a common locus. *Journal of Memory and Language*, *56*, 49–64. https://doi.org/10.1016/j.jml.2006.09.002

Patterson, K. E., Graham, N. L., & Hodges, J. R. (1994). The impact of semantic memory loss on phonological representations. *Journal of Cognitive Neuroscience*, *6*(1), 57–69. https://doi.org/10.1162/jocn.1994.6.1.57

Patterson, K. E., & Lambon Ralph, M. A. (1999). Selective disorders of reading? *Current Opinion in Neurobiology*, *9*(2), 235–239. https://doi.org/10.1016/S0959-4388(99)80033-6

Patterson, K. E., Lambon Ralph, M. A., Jefferies, E., Woollams, A. M., Jones, R., Hodges, J. R., & Rogers, T. T. (2006). “‘Presemantic’” Cognition in Semantic Dementia: Six Deficits in Search of an Explanation. *Journal of Cognitive Neuroscience*, *18*(2), 169–183. https://doi.org/10.1162/jocn.2006.18.2.169

Peters, F., Majerus, S., De Baerdemaeker, J., Salmon, E., & Collette, F. (2009). Impaired semantic knowledge underlies the reduced verbal short-term storage capacity in Alzheimer’s disease. *Neuropsychologia*, *47*(14), 3067–73. https://doi.org/10.1016/j.neuropsychologia.2009.07.002

Pettigrew, C., & Hillis, A. E. (2011). Role for Memory Capacity in Sentence Comprehension: Evidence from Acute Stroke. *Aphasiology*, *4*(164), 1258–1280. https://doi.org/10.1126/scisignal.2001449.Engineering

Plaut, D. C., & Kello, C. T. (1999). The emergence of phonology from the interplay of speech comprehension and production: A distributed connectionist approach. In B. MacWhinney (Ed.), *The emergence of language* (pp. 381–415). Mahweh, NJ: Erlbaum.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. E. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological Review*, *103*(1), 56–115. https://doi.org/10.1037/0033-295X.103.1.56

Plaut, D. C., & Shallice, T. (1993). Deep Dyslexia: A Case Study of Cormectionist Neuropsychology. *Cognitive Neuropsychology*, *10*(5), 377–500. https://doi.org/10.1080/02643299308253469

Poirier, M., & Saint-Aubin, J. (1995). Memory for related and unrelated words: further evidence on the influence of semantic factors in immediate serial recall. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *48*(2), 384–404. https://doi.org/10.1080/14640749508401396

Price, C. J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. *NeuroImage*, *62*(2), 816–847. https://doi.org/10.1016/j.neuroimage.2012.04.062

Protopapas, A. (2007). CheckVocal : A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods, 39*(4), 859–862. https://doi.org/10.3758/BF03192979

Rastle, K., Croot, K. P., Harrington, J. M., & Coltheart, M. (2005). Characterizing the motor execution stage of speech production: consonantal effects on delayed naming latency and onset duration. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(5), 1083–95. https://doi.org/10.1037/0096-1523.31.5.1083

Roberts, D. J., Lambon Ralph, M. A., & Woollams, A. M. (2010). When does less yield more? The impact of severity upon implicit recognition in pure alexia. *Neuropsychologia*, *48*(9), 2437–2446. https://doi.org/10.1016/j.neuropsychologia.2010.04.002

Rodd, J. M. (2004). The effect of semantic ambiguity on reading aloud: a twist in the tale. *Psychonomic Bulletin & Review*, *11*(3), 440–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15376792

Romani, C., McAlpine, S., & Martin, R. C. (2008). Concreteness effects in different tasks: Implications for models of short-term memory. *Quarterly Journal of Experimental Psychology*, *61*(2), 292–323. https://doi.org/10.1080/17470210601147747

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., … Weiller, C. (2008). Ventral and dorsal pathways for language. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(46), 18035–40. https://doi.org/10.1073/pnas.0805234105

Savill, N., Ellis, A. W., & Jefferies, E. (2017). Newly-acquired words are more phonologically robust in verbal short-term memory when they have associated semantic representations. *Neuropsychologia*, *98*, 85–97. https://doi.org/10.1016/j.neuropsychologia.2016.03.006

Savill, N., Ellis, R., Brooke, E., Koa, T., Ferguson, S., Rojas-Rodriguez, E., … Jefferies, E. (2018). Keeping it together: Semantic coherence stabilizes phonological sequences in short-term memory. *Memory and Cognition*, 1–12. https://doi.org/10.3758/s13421-017-0775-3

Savill, N., Metcalfe, T., Ellis, A. W., & Jefferies, E. (2015). Semantic categorisation of a word supports its phonological integrity in verbal short-term memory. *Journal of Memory and Language*, *84*, 128–138. https://doi.org/10.1016/j.jml.2015.06.003

Schock, J., Cortese, M. J., & Khanna, M. M. (2012). Imageability estimates for 3,000 disyllabic words. *Behavior Research Methods, 44*(2), 374–9. https://doi.org/10.3758/s13428-011-0162-0

Shibahara, N., Zorzi, M., Hill, M. P., Wydell, T., & Butterworth, B. (2003). Semantic effects in word naming: evidence from English and Japanese Kanji. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *56*(2), 263–286. https://doi.org/10.1080/02724980244000369

Strain, E., & Herdman, C. M. (1999). Imageability effects in word naming: an individual differences analysis. *Canadian Journal of Experimental Psychology*, *53*(4), 347–359. https://doi.org/10.1037/h0087322

Strain, E., Patterson, K. E., & Seidenberg, M. S. (1995). Semantic Effects in Single-Word Naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(5), 1140–1154. https://doi.org/10.1037/0278-7393.21.5.1140

Strain, E., Patterson, K. E., & Seidenberg, M. S. (2002). Theories of word naming interact with spelling--sound consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(1), 207–214. https://doi.org/10.1037//0278-7393.28.1.207

Taylor, J. S. H., Duff, F. J., Woollams, A. M., Monaghan, P., & Ricketts, J. (2015). How Word Meaning Influences Word Reading. *Current Directions in Psychological Science*, *24*(4), 322–328. https://doi.org/10.1177/0963721415574980

Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (1999). *Test of Word Reading Efficiency*. Austin, TX: PRO-ED Publishing, Inc.

Tse, C.-S., & Altarriba, J. (2007). Testing the associative-link hypothesis in immediate serial recall: Evidence from word frequency and word imageability effects. *Memory*, *15*(6), 675–90. https://doi.org/10.1080/09658210701467186

Tyler, L. K., Voice, J. K., & Moss, H. E. (2000). The interaction of meaning and sound in spoken word recognition. *Psychonomic Bulletin & Review*, *7*(2), 320–326. https://doi.org/10.3758/bf03212988

Ueno, T., Saito, S., Rogers, T. T., & Lambon Ralph, M. A. (2011). Lichtheim 2: Synthesizing aphasia and the neural basis of language in a neurocomputational model of the dual dorsal-ventral language pathways. *Neuron*, *72*(2), 385–396. https://doi.org/10.1016/j.neuron.2011.09.013

Ueno, T., Saito, S., Saito, A., Tanida, Y., Patterson, K. E., & Lambon Ralph, M. A. (2011). Not Lost in Translation : Generalization of the Primary Systems Hypothesis to Japanese-specific Language Processes. *Journal of Cognitive Neuroscience*, *26*(2), 433–446. https://doi.org/10.1162/jocn_a_00467

Vallar, G., & Baddeley, A. D. (1987). Phonological short-term store and sentence processing. *Cognitive Neuropsychology*, *4*(4), 417–438. https://doi.org/10.1080/02643298708252046

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Verhaegen, C., Piertot, F., & Poncelet, M. (2013). Dissociable components of phonological and lexical-semantic short-term memory and their relation to impaired word production in aphasia. *Cognitive Neuropsychology*, *30*(7–8), 544–563. https://doi.org/10.1080/02643294.2014.884058

Vitevitch, M. S. (2003). The influence of sublexical and lexical representations on the processing of spoken words in English. *Clinical Linguistics & Phonetics*, *17*(6), 487–499. https://doi.org/10.1016/j.immuni.2010.12.017.Two-stage

Vitevitch, M. S., Armbruster, J., & Chu, S. (2004). Sublexical and lexical representations in speech production: effects of phonotactic probability and onset density. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 514–29. https://doi.org/10.1037/0278-7393.30.2.514

Vitevitch, M. S., & Luce, P. A. (1998). WHEN WORDS COMPETE: Levels of Processing in

Perception of Spoken Words. *Psychological Science*, *9*(4), 325–329. https://doi.org/10.1111/1467-9280.00064

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*(3), 374–408. https://doi.org/10.1006/jmla.1998.2618

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 481–487. https://doi.org/10.3758/BF03195594

Vitevitch, M. S., & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, *52*(2), 193–204. https://doi.org/10.1016/j.jml.2004.10.003

Vitevitch, M. S., Luce, P. A., Pisoni, D. B., & Auer, E. T. (1999). Phonotactics, Neighborhood Activation, and Lexical Access for Spoken Words. *Brain and Language*, *68*(1–2), 306–311. https://doi.org/10.1124/dmd.107.016501.CYP3A4-Mediated

Walker, I., & Hulme, C. (1999). Concrete words are easier to recall than abstract words: Evidence for a semantic contribution to short-term serial recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(5), 1256–1271. https://doi.org/10.1037/0278-7393.25.5.1256

Warmington, M., Stothard, S. E., & Snowling, M. J. (2013). Assessing dyslexia in higher education: The York adult assessment battery-revised. *Journal of Research in Special Educational Needs*, *13*(1), 48–56. https://doi.org/10.1111/j.1471-3802.2012.01264.x

Warrington, E. K. (1997). The Graded Naming Test: A Restandardisation. *Neuropsychological Rehabilitation*, *7*(2), 143–146. https://doi.org/10.1080/713755528

Warrington, E. K., McKenna, P., & Orpwood, L. (1998). Single Word Comprehension: A Concrete and Abstract Word Synonym Test. *Neuropsychological Rehabilitation*, *8*(2), 143–154. https://doi.org/10.1080/713755564

Welbourne, S. R., Woollams, A. M., Crisp, J., & Lambon Ralph, M. A. (2011). The role of plasticity-related functional reorganization in the explanation of central dyslexias. *Cognitive Neuropsychology*, *28*(2), 65–108. https://doi.org/10.1080/02643294.2011.621937

Wilshire, C. E., & Fisher, C. (2004). “Phonological” Dysphasia: a Cross-Modal Phonological Impairment Affecting Repetition, Production, and Comprehension. *Cognitive Neuropsychology*, *21*(2), 187–210. https://doi.org/10.1080/02643290342000555

Wilshire, C. E., Keall, L. M., & O’Donnell, D. J. (2010). Semantic contributions to immediate serial recall: evidence from two contrasting aphasic individuals. *Neurocase*, *16*(4), 331–351. https://doi.org/10.1080/13554791003620256

Woollams, A. M. (2005). Imageability and ambiguity effects in speeded naming: convergence and divergence. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *31*(5), 878–90. https://doi.org/10.1037/0278-7393.31.5.878

Woollams, A. M. (2015). For richer or poorer? Imageability effects in semantic dementia patients’ reading aloud. *Neuropsychologia*, *76*, 254–263. https://doi.org/10.1016/j.neuropsychologia.2015.03.023

Woollams, A. M., Hoffman, P., Roberts, D. J., Lambon Ralph, M. A., & Patterson, K. E. (2014). What lies beneath: A comparison of reading aloud in pure alexia and semantic dementia. *Cognitive Neuropsychology*, *31*(5–6), 37–41. https://doi.org/10.1080/02643294.2014.882300

Woollams, A. M., Lambon Ralph, M. A., Madrid, G., & Patterson, K. E. (2016). Do You Read How I Read? Systematic Individual Differences in Semantic Reliance amongst Normal Readers. *Frontiers in Psychology*, *7*, 1757. https://doi.org/10.3389/fpsyg.2016.01757

Woollams, A. M., Lambon Ralph, M. A., Plaut, D. C., & Patterson, K. E. (2007). SD-squared: On the association between semantic dementia and surface dyslexia. *Psychological Review*, *114*(2), 316–339. https://doi.org/10.1037/0033-295X.114.2.316

Figure Legends

Figure 1. Analysis of the effects of imageability and inverse age-of-acquisition (AoA) on response times in reading and repetition and on accuracy in immediate serial recall. Reading, Rep and ISR: Reading, single-word repetition and immediate serial recall respectively. Y-axis shows modelled mean performance. X-axis shows z-scored imageability rating (left-hand column) and inverse AoA rating (right-hand column). There were effects of both of these markers of lexical-semantic knowledge in all tasks – e.g., faster reading and repetition and more accurate ISR for words with higher imageability or inverse AoA. For reading, the red line shows items with consistent spelling-sound correspondences, while the blue line shows inconsistent items. For repetition, the red and blue lines depict performance for items with high and low onset phonotactic probabilities respectively.

Figure 2. The relationship between word and nonword performance on each task, plotted separately for high and low imageability items. Blue line: low imageability items (associated with slower RT and lower accuracy). Red line: high imageability items. Black line: Expected performance on high imageability items under the null hypothesis – i.e., with no relationship between the effect of imageability and nonword performance.

Figure 3. Correlations between the effect of imageability and nonword performance across tasks.

Figure 4. Partial correlations between the effects of imageability across tasks and phonological and semantic performance.

Table 1. *Average stimulus properties for each experimental condition in the immediate serial recall task*

	High imageability words	Low imageability words	Nonwords
Imageability Rating (1-7)	5.84 (0.66)	2.83 (0.54)	N/A
Age of Acquisition	7.10 (2.19)	9.59 (2.56)	N/A
Frequency	3.50 (0.38)	3.49 (0.68)	N/A
Letters	4.22 (0.73)	4.18 (0.75)	N/A
Syllables	1.00 (0.00)	1.00 (0.00)	1.00 (0.00)
Phonemes	3.00 (0.00)	3.00 (0.00)	3.00 (0.00)
Summed Biphone Prob.	.006 (0.006)	.006 (0.005)	.005 (0.007)

Note. Average properties of the immediate serial recall test items are shown with standard deviations in parentheses.

Table 2. *Average stimulus properties for each experimental condition in the single-item reading task*

	High imageability words		Low imageability words		Nonwords	
	Inconsistent “bruise”	Regular “chimp”	Inconsistent “guise”	Regular “stint”	Inconsistent “bouge”	Regular “larse”
Imageability Rating (1-7)	5.92 (0.59)		2.66 (0.50)		N/A	
	5.87 (0.61)	5.97 (0.56)	2.63 (0.53)	2.69 (0.44)		
Age of Acquisition	7.86 (2.25)		11.14 (2.08)		N/A	
	7.82 (2.25)	7.91 (2.27)	11.62 (2.11)	10.67 (1.6)		
Frequency	3.28 (0.41)		3.19 (0.52)		N/A	
	3.32 (0.47)	3.23(0.33)	3.21 (0.51)	3.18 (0.54)		
Letters	5.19 (1.07)		5.24 (1.06)		5.33 (0.91)	
	5.22 (1.09)	5.17 (1.06)	5.30 (1.12)	5.18 (1.00)	5.20 (0.89)	5.30 (0.88)
Syllables	1.33 (0.47)		1.33 (0.47)		1.33 (0.47)	
	1.33 (0.48)	1.33 (0.48)	1.33 (0.48)	1.33 (0.48)	1.33 (0.48)	1.33 (0.48)
Phonemes	3.93 (0.99)		3.89 (1.04)		3.95 (1.00)	
	3.82 (0.91)	4.03 (1.05)	3.80 (1.00)	4.02 (1.08)	3.86 (0.95)	4.05 (1.04)
Summed Biphone Prob.	.012 (0.010)		.012 (0.011)		.011 (0.010)	
	.009 (0.008)	.015 (0.012)	.010 (0.009)	.014 (0.012)	.008 (0.006)	.014 (0.011)
Execution-Acoustic Interval	264.11 (43.58)		262.69 (37.34)		268.72 (35.86)	
	263.75 (55.84)	264.47 (26.68)	266.12 (28.25)	259.32 (44.50)	270.47 (43.32)	267.53 (26.97)

Note. An example item from each reading consistency condition is shown in quotation marks. The corresponding average properties are given beneath the collapsed averages corresponding to overall imageability condition. Standard deviations are in parentheses.

Table 3. *Average stimulus properties for each experimental condition in the single-item repetition task*

	High imageability words		Low imageability words		Nonwords	
	High Phon. “wrist”	Low Phon. “thong”	High Phon. “cusp”	Low Phon. “poise”	High Phon. “carge”	Low Phon. “sheng”
Imageability Rating (1-7)	6.10 (0.62)		2.68 (0.49)		N/A	
	6.11 (0.62)	6.09 (0.61)	2.69 (0.43)	2.68 (0.55)		
Age of Acquisition	6.95 (2.24)		10.50 (2.46)		N/A	
	7.07 (1.98)	6.83 (2.49)	10.39 (2.15)	10.61 (2.74)		
Frequency	3.59 (0.47)		3.51 (0.63)		N/A	
	3.59 (0.43)	3.58 (0.51)	3.54 (0.64)	3.47 (0.64)		
Letters	5.08 (1.25)		5.13 (1.04)		N/A	
	4.93 (1.27)	5.23 (1.23)	4.93 (1.00)	5.32 (1.07)		
Syllables	1.33 (0.47)		1.33 (0.47)		1.33 (0.47)	
	1.33 (0.48)	1.33 (0.48)	1.33 (0.48)	1.33 (0.48)	1.33 (0.48)	1.33 (0.48)
Phonemes	3.88 (1.03)		3.93 (1.09)		3.96 (1.10)	
	3.87 (1.07)	3.88 (0.99)	3.95 (1.11)	3.92 (1.09)	3.98 (1.08)	3.93 (1.12)
Summed Biphone Prob.	.013 (0.011)		.015 (0.012)		.012 (0.013)	
	.017 (0.013)	.008 (0.008)	.019 (0.013)	.011 (0.010)	.020 (0.014)	.005 (0.004)
Execution-Acoustic Interval	268.36 (24.55)		267.69 (25.58)		268.99 (25.37)	
	269.88 (23.25)	266.82 (25.93)	266.87 (23.75)	268.46 (27.40)	273.57 (23.45)	264.57 (26.57)

Note. An example item from each phonotactic condition (based on onset CV/VC) is shown in quotation marks. The corresponding average properties are given beneath the collapsed averages corresponding to overall imageability condition. Standard deviations are in parentheses.

Table 4. *Effects of imageability and age-of-acquisition (AoA), examined in separate models, for reading, single-item repetition and immediate serial recall tasks.*

Model Parameter	**F-value (DF)**	**Z-value**	**p-value**	**Estimate**	**95% CI**	**-2Log likelihood**
Reading						
Empty model						208005.5
Full model (imageability)						
Fixed effects:						
Stimulus exposure	5.24(2, 15000)		.0053	1: 9.55	3.37 – 15.74	198079.4
				2: 1.95	-4.42 – 8.33	
Imageability	15.04 (1, 234)		.0001	-18.57	-26.67 – -10.47	198061.7
Consistency	47.74 (1, 228)		<.0001	-39.44	-50.62 – -28.25	198013.5
Consistency × Imageability	6.60 (1, 229)		.011	14.72	3.49 – 25.95	198001.7
Random effects:						
Ppt variance: intercept		6.35	<.0001	4538.74		
Ppt variance: slope (for imageability)		1.70	.045	12.31		
Ppt covariance: intercept and slope		-3.15	.0016	-172.32		
Item variance: intercept		10.09	<.0001	181.25		
Full model (1/AoA)						
Fixed effects:						
Stimulus exposure	4.87 (2, 14000)		.0077	1: 9.13	3.02 – 15.25	198079.4
				2: 1.98	-4.37 – 8.32	
1/AoA	48.87 (1, 257)		<.0001	-28.67	-35.98 – -21.35	198029.0
Consistency	53.39 (1, 228)		<.0001	-38.37	-48.66 – -28.07	197977.9
Consistency × 1/AoA	12.32 (1, 229)		.0004	18.59	8.21 – 28.97	197960.8
Random effects:						
Ppt variance: intercept		6.35	<.0001	4596.44		
Ppt variance: slope (for 1/AoA)		3.97	<.0001	55.98		
Ppt covariance: intercept and slope		-4.49	<.0001	-371.16		
Item variance: intercept		10.01	<.0001	1530.61		
Repetition						
Empty model						222397.3
Full model (imageability)						
Fixed effects:						
Block	8.70 (3, 239)		<.0001	1: 39.47	23.27 – 55.67	222372.4
				2: 22.38	6.18 – 38.58	
				3: 8.62	-7.58 – 24.82	
Stimulus exposure	3.30 (2, 16000)		.037	1: 4.36	-2.37 –11.09	222366.1
				2: -4.56	-11.63 – 2.51	
Imageability	6.13 (1, 243)		.015	-7.25	-12.99 – -1.51	222360.0
Random effects:						
Ppt variance: intercept		6.43	<.0001	14336		
Item variance: intercept		10.32	<.0001	1932.8		
Full model (1/AoA)						
Fixed effects:						
Block	8.13 (3, 239)		<.0001	1: 36.90	21.02 – 52.78	222372.4
				2: 21.51	5.68 – 37.34	
				3: 7.14	-8.69 – 22.97	
Stimulus exposure	3.63 (2, 16000)		.026	1: 5.19	-1.56 –11.94	222366.1
				2: -4.06	-11.16 – 3.05	
1/AoA	16.52 (1, 241)		<.0001	-12.12	-18.05 – -6.19	222348.9
Random effects:						
Ppt variance: intercept		6.43	<.0001	14336		
Ppt variance: slope (for 1/AoA)		3.60	.0002	51.82		
Item variance: intercept		10.30	<.0001	1838.77		

Immediate Serial Recall						
Empty model						99110.44
Full model (imageability)						
Fixed effects:						
Serial position	55.42 (1, 216)		<.0001	-0.017	-0.021 – -0.013	99001.41
List	14.53 (1, 215.6)		.0002	-0.012	-0.018 – -0.0058	98994.16
Imageability	10.22 (1, 215.6)		.0016	0.21	0.081 – 0.34	98995.92
Random effects:						
Ppt variance: intercept		5.95	<.0001	0.29		
Item variance: intercept		9.71	<.0001	1.02		
Full model (1/AoA)						
Fixed effects:						
Serial position	54.62 (1, 215.7)		<.0001	-0.017	-0.021 – -0.012	99001.41
List	15.96 (1, 215.2)		<.0001	-0.012	-0.019 – -0.0064	98994.16
1/AoA	11.90 (1, 215.7)		.0007	0.23	0.097 – 0.35	98988.40
Random effects:						
Ppt variance: intercept		5.95	<.0001	0.29		
Item variance: intercept		9.71	<.0001	1.02		

The table lists effects that were included in the final model (on the basis that they significantly improved model fit). A full description of the variables we considered is provided in the text. In the repetition task, phonotactic probability and its interaction with imageability did not significantly improve model fit and were not included in the final model. Imageability = Word imageability rating. 1/AoA = 1/Age of Acquisition word rating (inverse score used for comparability with imageability). Consistency = Spelling-to-sound consistency (consistent, inconsistent). Stimulus exposure = Captures whether the first, second or third presentation of the stimulus across tasks, for a given individual (‘core’ stimuli were presented in each task; other stimuli were presented once only). Block = Presentation block number within the reading or repetition task. Serial position = serial position of item within an ISR list. List = ISR list identifier. For each model, the -2Log-Likelihood values start with the empty model in the first row. Successive rows show the reduced -2Log-likelihood having added additional fixed effects. The last value corresponds to the -2Log-likelihood for the full model. All other parameter, variance and covariance estimates are derived from the full model.

Figure 1

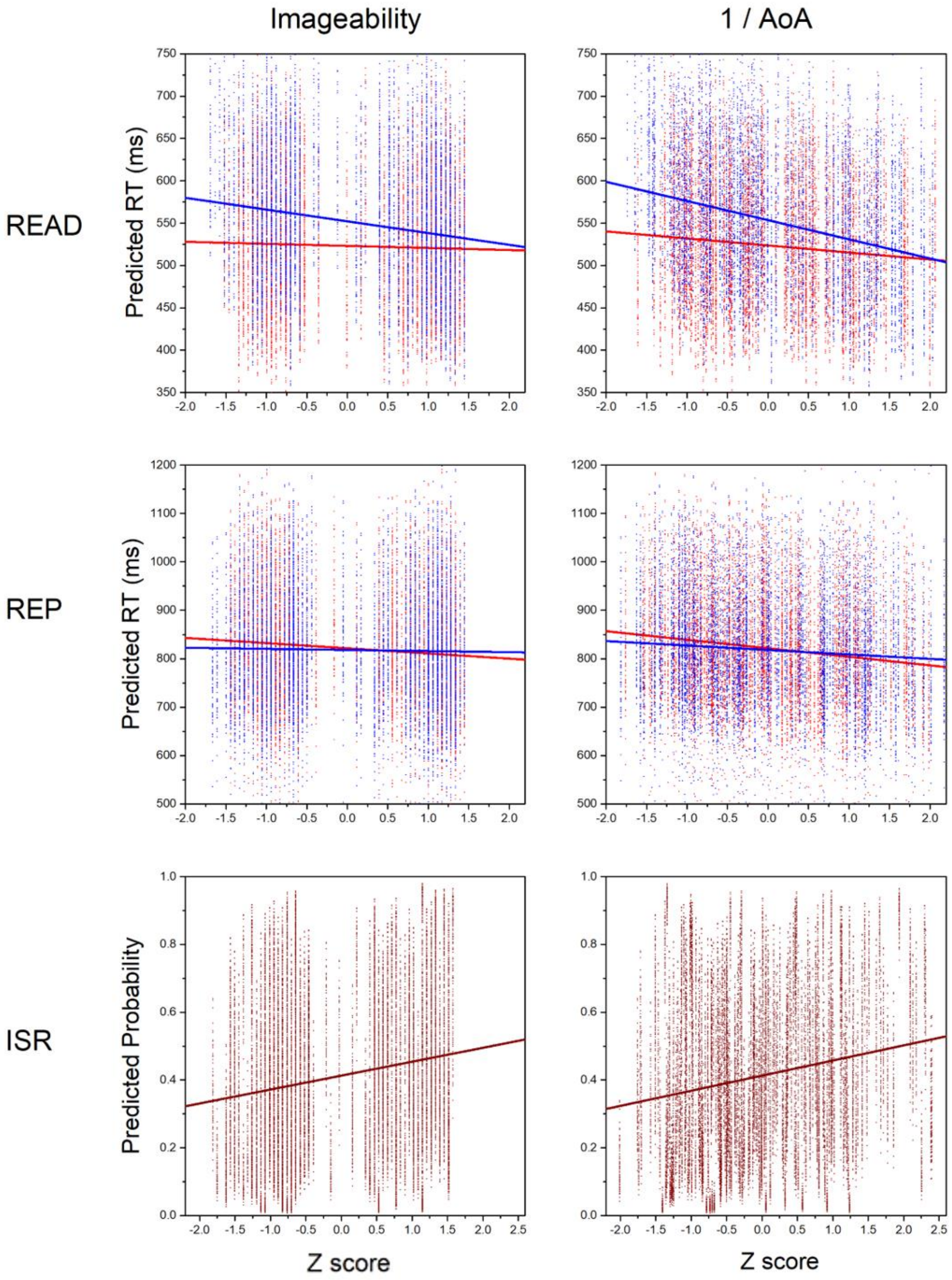
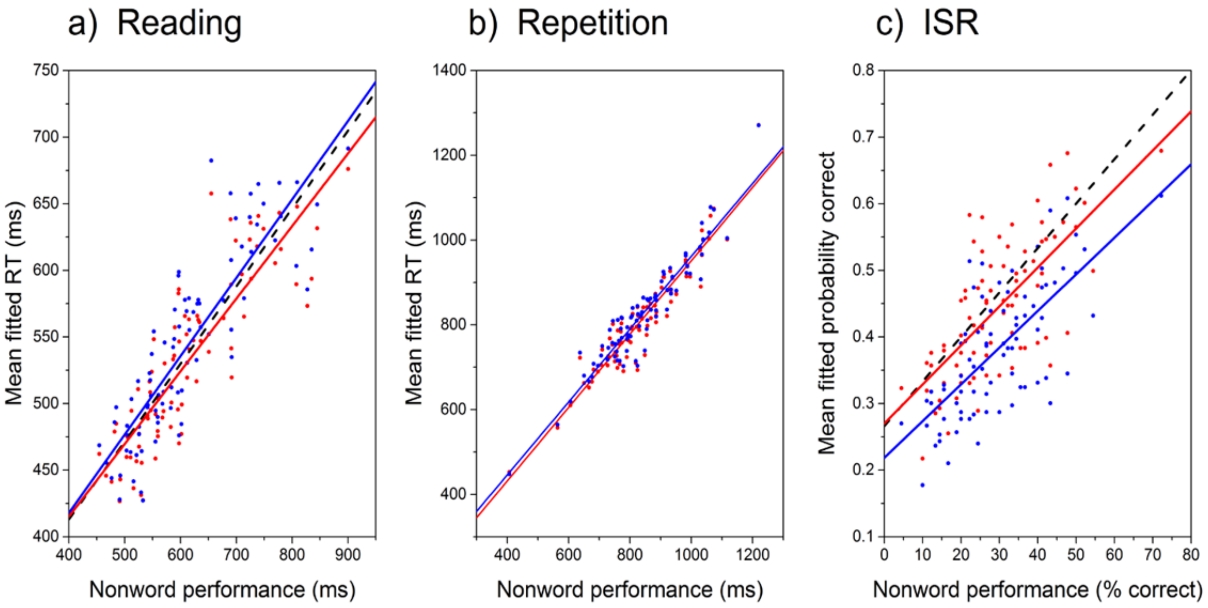
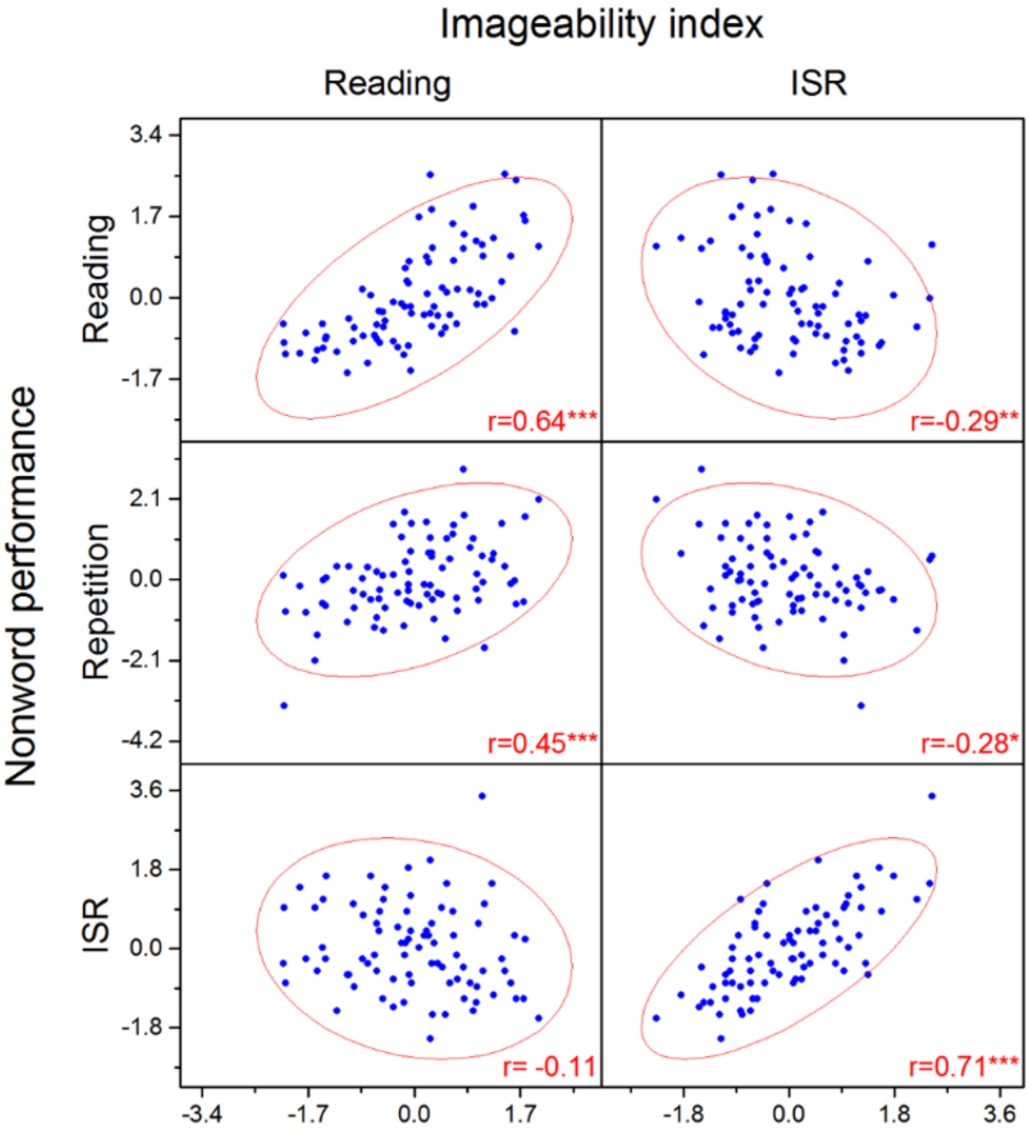

Imageability
1 / AoA
READ
REP
ISR
Predicted RT (ms)
Predicted RT (ms)
Predicted Probability
750
700
650
600
550
500
450
400
350
1200
1100
1000
900
800
700
600
500
1.0
0.8
0.6
0.4
0.2
0.0
-2.0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
Z score
Z score

Figure 2



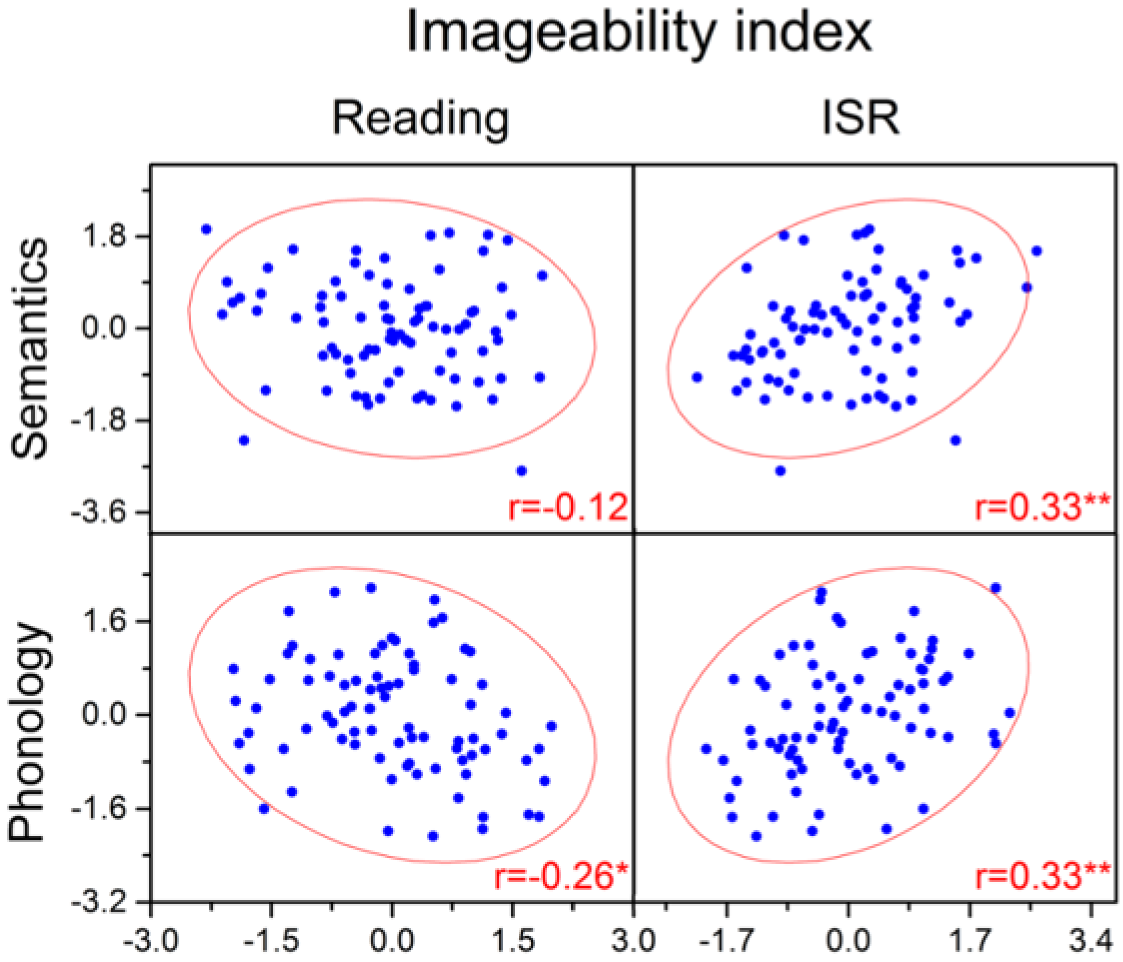
a) Reading
Mean fitted RT (ms)
750
700
650
600
550
500
450
400
400
500
600
700
800
900
Nonword performance (ms)
b) Repetition
Mean fitted RT (ms)
1400
1200
1000
800
600
400
400
600
800
1000
1200
Nonword performance (ms)
c) ISR
Mean fitted probability correct
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
10
20
30
40
50
60
70
80
Nonword performance (% correct)

Figure 3



NB * = p<.05, ** = p<.01, *** = p<.0001

Figure 4



NB * = p<.05, ** = p<.01, *** = p<.0001