【Research Article】

To What Extent Does Form-Based Priming Account for the Mnemonic Effect of Phonological Patterns? A Preliminary Investigation

Michael Green

Abstract

There is a growing body of evidence that suggests that phonological patterns in multi-word units may act as mnemonic devices and thus aid in the acquisition of L2 vocabulary. Some researchers state that learners' attention should be deliberately drawn to such patterns, though there are conflicting findings in the literature on this point. One explanation that has been posited for the mnemonic influence involves a form-based priming effect. If it can be shown that phonological patterns automatically aid lexical processing, the intervention to raise awareness of such patterns may not be necessary. To shed light on this issue, the experiment reported herein is a preliminary investigation into the extent to which form-based priming accounts for the mnemonic effects observed in the literature. A lexical decision task (LDT) was conducted with 24 English L1 speakers to ascertain if alliteration and assonance facilitate lexical processing. The data from the LDT do not support any strong claims that perceptual priming is the determining factor for the processing advantage, which indicates the need for explicit attention to the patterns. Limitations and further avenues for investigation are also discussed.

Keywords: *Form-based priming, Phonological patterns, Vocabulary acquisition*

Introduction

It seems almost axiomatic that phonological patterns such as rhyme, alliteration[1], and assonance[2], can act as aids to memory; for instance, rhyme can be used to learn historical events (*In fourteen hundred and ninety-two, Columbus sailed the ocean blue*) or the number of days in the month (*Thirty days hath September, April, June and November…*). There is also a longstanding and widespread use of phonological patterns in advertising, slogans, and brand names (*Coca Cola*, *Kit Kat*, and *Tim Tams*), fictional characters (*Donald Duck* and *Wonder Woman*), nursery rhymes, and of course, poetry and prose. Research by Jusczyk et

1 Operationalized herein as the repetition, in two consecutive words, of the same consonant sound in an initial stressed syllable, e.g., *tall tree*.

2 The repetition of a vowel or diphthong sound in the prominent syllable of two consecutive words, e.g., *main gate.*

al. (1999) shows that even nine-month-old infants are already sensitive to alliteration in consonant-vowel-consonant syllables, which suggests that patterns of similarity are pertinent to building L1 vocabulary. Phonological patterns also emerge in early childhood wordplay, reflecting the ludic or playful function of language (Crystal, 1997), and in Peters' (1983) seminal account of first language acquisition, phonological patterns are posited as heuristic devices that young children employ to segment or compare "units" of meaning in the intermittent stream of speech. The literature on oral traditions and the genres of epics and ballads (for a comprehensive survey, see Rubin, 1995) proposes that phonological patterns, together with meaning and imagery, are constraints that cue memories and restrict choices, enabling thousands of lines of songs, stories, and poems to be memorized and transmitted for centuries.

In the field of second language acquisition, Nation (2001, 2014) notes that English multi-word units (MWUs), such as binomials, collocations, compounds, and idioms, often alliterate or assonate and advises L2 learners to pay deliberate attention to these patterns, advice that has been repeated more recently by Szudarski (2017). This advice is partially predicated on empirical work from a single group of researchers (Table 1). This series of small-scale, classroom-based quasi-experiments all explore whether English L2 learners can identify or recall MWUs with phonological patterns better than equivalent[3] sequences with no such pattern. In general, the experiments involve teacher-led dictations of target MWUs followed by a series of free- and cued-recall tests over various time intervals (immediate tests and delayed tests up to two weeks). The designs are generally iterative, involving a partial replication of a previous experiment, and collectively, these studies build the case that phonological patterns in MWUs have a mnemonic effect on L2 learners to some degree.

Table 1

An Overview of the Findings in the Experiments of Boers et al. on Phonological Patterns

	Study (*phonological pattern*)		**Awareness-Raising?**	**Statistically Significant Finding?**
1	Lindstromberg & Boers (2008a) *(alliteration)*	Exp 1	Yes	Yes
		Exp 3	Yes	
2	Lindstromberg & Boers (2008b) *(assonance)*		Yes	Yes
3	Boers et al. (2012) *(alliteration)*			Yes
4	Boers et al. (2014a) *(alliteration)*	Exp 1		Yes
		Exp 2		Yes
5	Boers et al. (2014b) *(assonance)*	Exp 1		
		Exp 2	Yes	Yes
6	Boers et al. (2014c) *(consonance)*			Yes [1]
7	Boers, Eyckmans & Lindstromberg (2014) *(alliteration and consonance)*		Yes	

3 Experimental items are often balanced in terms of length, frequency, collocational strength (with reference to MI or t-scores), concreteness and imageability, and L1 cognate status.

	Study (*phonological pattern*)		Awareness-Raising?	Statistically Significant Finding?
8	Lindstromberg & Eyckmans (2014) *(assonance)*	Exp 1	Yes	Yes
		Exp 2	Yes	Yes
9	Boers et al. (2014) *(alliteration)*			Yes
10	Eyckmans et al. (2016) *(alliteration)*		Yes	Yes
11	Eyckmans & Lindstromberg (2017) *(alliteration and assonance)*		Yes	Yes
12	Lindstromberg & Eyckmans (2017) (*assonance*)		Yes	

[1] A statistically significant finding, but not in the predicted direction.

An issue that arises is whether participants need to have the phonological pattern highlighted to benefit from any mnemonic effect. As can be seen from the *Awareness-Raising* column in Table 1, there is inconsistent evidence on this point; statistically significant results have been found when there is no attention-raising (e.g., Studies 3, 4, and 9), and non-significant results found when the patterns have been brought to the attention of the participants (e.g., Studies 1(3), 7, and 12).

One possible explanation for the facilitation of lexical retrieval of MWUs displaying phonological patterns is the short-term memory process of priming, an explanation that the researchers have posited in the past (e.g., Boers et al., 2014a; Eyckmans et al., 2016; Eyckmans & Lindstromberg, 2017; Lindstromberg & Boers, 2008a; Lindstromberg & Eyckmans, 2017). If priming is the principal mechanism behind which the mnemonic effect operates, no teacher intervention would be necessary in theory as priming is held to be an implicit memory process and is thus often described as unintentional or unconscious (Tulving & Schacter, 1990). To further tease apart this issue and to see if an awareness-raising component is indeed warranted, a lexical decision task (LDT) was performed to assess the plausibility of priming as an explanatory factor for the mnemonic advantage of alliterating or assonating MWUs, as seen in the experiments outlined in Table 1.

Literature Review

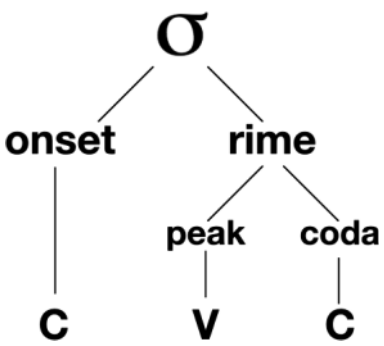
The term “priming” appears to do a lot of heavy lifting in the literature: it can refer to a *component* of implicit long-term memory (Baddeley, 2001) or a research *paradigm* or *tool* for examining lexical processing (McDonough & Trofimovich, 2009). Moreover, Shallice and Cooper (2011) variously employ the term “priming” to cover a functional imaging *procedure*, a short-term memory *process*, a general *characteristic* of cognitive subsystems, and a *property* of processing. Distinctions between these disparate usages are not always clearly articulated. Such terminological diffusion may have arisen partly because priming has been the focus of research across several disciplines, notably neuropsychology, cognitive psychology, psycholinguistics, and, increasingly, second language acquisition studies.

Broadly speaking, “priming” is used here as a term to describe how prior experience with language can unconsciously influence the processing of subsequent language. One reason phonological patterns such as alliteration and assonance are thought to be mnemonic devices is that form-based aspects of a given word privilege the processing of a following word that shares the same form-based aspects, whether they are phonological, orthographic, or phonetic.

The literature on form or perceptual priming is extensive, so when looking for evidence that alliteration or assonance can aid lexical processing, it may help first determine how such phonological patterns are operationalized within the priming paradigm. In effect, the question of interest is whether hearing a word beginning with or containing a particular sound will make it easier to process a subsequent word with the same sound in the same position. The form-based similarity of a prime and its target is frequently manipulated in terms of matching or overlapping particular segments of the two words and/or manipulating the number of segments that match. This segmentation is often based on a model of syllabic division (Figure 1).

Figure 1

Segmentation of a Consonant-Vowel-Consonant Monosyllable



Thus, a facilitatory priming effect for alliteration would involve test items with matching phonological consonantal onsets and no other relationship—morphological, syntactic, or semantic—as they could themselves account for any observed priming effect. Likewise, a priming effect for assonance would be based only on the single aspect of a matching stressed vowel sound (the peak or nucleus).

However, in Zwitserlood's (1996) overview of the form priming experimental paradigm, the stimuli overlap is described as "word-initial or rhyme" (p. 590), with no scope for a priming effect based on assonance alone. Rhyme is usually operationalized as phonological similarity of both the syllabic peak and coda (the rime), or in other words, similarity starting from the stressed vowel sound to the word offset. As we shall see below, perceptual priming based on assonance alone has received very little attention. As the findings for word-initial similarity are inconsistent, those based on rhyme will be considered first. Indeed, much of the most robust evidence is found in studies where the prime-target relationship is that of rhyme. Thus, participants will identify the target letter string *BEAN* more accurately and faster after being exposed to the prime *mean* than after the prime *pink*. This systematic priming effect for rhyme has been found in both monosyllabic and bisyllabic prime-target pairs, irrespective of frequency, for both words and pseudowords, and independently of the task performed: LDTs (Norris et al., 2002), shadowing (Dumay et al., 2001), and identification in noise (Slowiaczek et al., 1987).

In contrast to the evidence that rhyme can prime, even under different experimental designs, the situation for a priming effect for word-initial similarity is more complex and inconsistent, with facilitatory, inhibitory, and null effects being reported. Although the same dependent variable of response latency is usually the focus across experimental designs, task differences can introduce different variables, so it may be helpful to focus on just one task, the LDT, to glean insight from the results. When looking at word-initial overlap that forms an alliterative relationship between the prime and the target, a set of "rather messy results" (Dumay et al., 2001, p. 121) can be found in the literature, as seen in five representative studies in Table 2.

Table 2

Examples of LDT Studies that Include Alliterative Word-Initial Overlap

Study	Examples of *primes: TARGETS*	Effect [1]
1 Slowiaczek & Pisoni (1986) (Experiment 1)	*black, bland, bleed, burnt, /blæt/, /blim/, /brɛm/: BLACK*	No effect [2]
2 Radeau et al. (1989) (Experiment 2)	*palais, poulet, rouler: PARURE* (French L1 speakers)	No effect [2]
3 Goldinger et al. (1992) (Experiment 3)	*bang: BONE*	Facilitation [3]
4 Monsell & Hirsh (1998) (Experiment 1)	*broom: BRUISE*	Inhibition
5 McQueen & Sereno (2005)	*zeep: ZOON* (Dutch L1 speakers)	No effect

[1] Slowiaczek and Hamburger (1992) propose facilitation is due to activation/excitation at a prelexical phoneme level, and inhibition is the result of competition between words at the lexical level.

[2] A facilitatory effect, in terms of faster reaction times, was found only when the prime and the target were identical (a repetition effect).

[3] An auditory priming technique was used, and a facilitatory effect was found only when the targets were presented in white noise.

From the examples given in Table 2, it is clear that the nature of some of the prime-target relationships goes beyond a simple alliterative pattern and encompasses a matching vowel nucleus (e.g., broom – BRUISE in Monsell & Hirsh, 1998).

In Studies 1 and 2, no priming effect was found when primes shared one, two, or three phonemes with the target. The use of real words as primes further confounds the issue as the relationship between prime and target is no longer purely form-based: for example, in Study 1, if *burnt* had shown a priming effect for *BLACK*, it might have been due to semantic association rather than form-based similarity.[4]

It might seem that the best way to avoid semantic priming would be to use pseudowords. However, these too can create difficulties, and the construction of the pseudowords may also be a factor in the lack of clear results. For example, in Study 1, some of the pseudoword targets are pseudohomophones (a nonword that sounds like a real word), such as /*stik*/ /*skot*/ /*bæns*/ and /*slæk*/. Jiang (2012) advises against the use of such items as several studies have shown that pseudohomophones generally take longer to reject than pseudowords, and these could affect the mean response times (RTs) in the different conditions.

The experiments in Table 2 also incorporated different timing elements in the designs, specifically the interstimulus intervals, the interval between the offset of the prime and the onset of the target, and differences in stimulus onset asynchrony, the interval between the onset of the prime and the onset of the target. Lengthening these two variables is often claimed to increase strategy use (Radeau et al., 1989). Although none of these was itself uncontrolled within its study, they are all potential variables determining the opportunity for an observable effect, and between them, these variations dilute the strength of evidence for initial consonant priming.

The failure to find a facilitatory effect with reaction time measures could be due to differences in methodologies, dependent measures, and/or stimuli, and as reiterated by Goldinger et al. (1992), null results do not support any definitive conclusions. Rastle and Brysbaert's (2006) exhaustive meta-analysis of the relevant priming literature concluded that phonological primes *do* have an effect, and some experimental results cannot be attributed to task-based biases or strategy use. However, the evidence for a priming effect based solely on alliteration is ambiguous. Furthermore, as mentioned previously, there is scant data to make any reliable assumptions about an equivalent effect for an assonating pattern.

Methodology

Experiment Design

One challenge facing the researcher is that although priming experiments have a long history in some form or another, "research labs accumulate a lot of informal knowledge about how to run particular [priming] experiments, which is rarely published" (Rastle & Brysbaert, 2006, p. 185). As a result, it can be difficult to ascertain exactly how such experiments are designed and implemented from previously published studies alone. Fortunately, for one of the standard methods used in priming research, namely reaction time studies, there exists a body of more prescriptive work with more explicit guidelines on the protocols (Jiang, 2012; McDonough & Trofimovich, 2009). Though there are a wide range of experimental tasks for measuring and quantifying priming effects, the most widely used for investigating the mental lexicon is the LDT. In a standard LDT, two stimuli are presented successively on a computer screen, first the "prime" and then the "target." The experimental task requires the participant to respond only to the target by either pressing the "Yes" button on a keyboard if they think

4　Aside from the *burnt-BLACK* example given in the paper, the authors only list the target items in the appendix, not the items used as primes; hence, it is not possible to ascertain the overall number of potential semantic associations between the test items.

the target is a real word or the “No” button if they decide the target is not a real word. Participants are asked to respond as fast and as accurately as possible to each target. The reaction times, mediated by the participant’s motor responses, are thought to provide indirect evidence for underlying cognitive processes. Priming is said to occur if the prime facilitates the response to the target in terms of faster reaction times (measured in milliseconds) and increased accuracy.

Facilitation can occur because of a conceptual or semantic relationship between the prime and the target: for example, participants react faster to the visual string *NURSE* if they have just been presented with *doctor*. However, as this experiment seeks to establish if there is a facilitatory effect due solely to a perceptual or form-based relationship between the prime and the target, semantic variables need to be controlled for (discussed presently).

As priming is thought to be an implicit process that precludes the use of strategies in language processing, the initial visual stimuli can be hidden or “masked.” In the priming research paradigm, masking is often accomplished in two ways. Firstly, the prime stimulus is presented so briefly on screen (e.g., 50–60 ms) that the participant is unaware of it. Secondly, a series of characters, often a string of hash keys as long as the target string (######), can be used as a forward mask preceding the prime, and the target word-string used as a backward mask (Jiang, 2012). These masking techniques are thought to block retinal after-images and pixel overlap, overwriting any visuo-sensory representations. Such steps help ensure that the participants cannot employ conscious strategies such as attempting to guess upcoming targets, trying to find the relationship between the primes and targets, or tuning in to experimental “bias” such as word length for primes and targets.

Before outlining the procedures adopted for this study, a further challenge needs to be addressed. There is growing empirical evidence that the bilingual lexicon is non-selective (Brysbaert, 2003; Nakayama et al., 2012). This premise means that the automatic activation of a phonological representation by a visual word stimulus is not limited to a representation specific to the language being read. This leads to the conclusion that if a participant has knowledge of more than one language—and if, as supposed, the first stages of word recognition are indeed language independent—it is possible to prime a target word in the L2 by a homophonic stimulus of the L1. There is an increasing amount of support for this notion in the literature; for example, for French-Dutch bilinguals, a French target (such as *OUI*) can be primed by a phonologically similar Dutch word (*wie*) (example from Experiment 1 in Brysbaert et al., 1999). There is also evidence of cross-language cognate priming across disparate writing systems. For example, Nakayama et al. (2012) found significant priming effects with Japanese-English bilinguals, where a masked stimuli such as ガイド /$_{\mu}$ga $_{\mu}$ɪ $_{\mu}$dɒ/ (guide) primed the English target *GUIDE*.

These findings raise four intriguing questions. Firstly, is this priming effect uni- or bi-directional? That is, would the L2 stimulus *guide* prime the L1 target ガイド (/$_{\mu}$ga $_{\mu}$ɪ $_{\mu}$dɒ/) for Japanese-English bilinguals? Secondly, what is the role of proficiency? Are the same priming effects found with less-proficient bilinguals such as adult L2 learners? Thirdly, does the priming effect only apply to alphabets and syllabaries or also to logographic scripts such as Japanese Kanji? Finally, is there a cross-language priming effect when the primes and targets are not cognates and therefore have no semantic relationship? For example, with

Japanese-English bilinguals would the L2 prime *guy* facilitate the processing of the L1 ガイド /$_{\mu}$ga $_{\mu}$ɪ $_{\mu}$dɒ/ and vice versa?

These issues would seem to suggest that it is thus paramount to rigorously control for pseudohomophone and cognate status in a cross-language LDT. However, numerous other lexical characteristics can also affect the processing of stimuli in such tasks and thus affect the participant's reaction time and accuracy. According to Jiang (2012), the following properties have all been found to affect lexical processing:

> frequency, familiarity, word length (in terms of letter, syllables, phonemes or morphemes), neighbourhood density, neighbourhood frequency, concreteness, imageability, age of acquisition, spelling-sound regularity, affixation, polysemy, bigram frequency, number of associates, lexicality, nonword legality, pseudohomophone and cognate status (pp. 80–82).

A further challenge is that inconsistent results are not uncommon in reaction time studies, and the exact effect of these variables is not without controversy (Jiang, 2012). Furthermore, different theoretical positions can complexify the explanation of apparently simple effects. For example, word frequency is often considered one of the most robust predictors of word recognition performance. It has long been known that participants respond more quickly and accurately to high-frequency than low-frequency words across virtually all lexical processing tasks (Whaley, 1978, cited in Yap & Balota, 2015). Although there are various potential models that explain the word frequency effect, most reflect a version of frequency affecting baseline activation. For instance, according to the interactive activation model (McClelland & Rumelhart, 1981), high-frequency words are responded to faster because they have a higher resting activation level (or lower threshold) and thus require less stimulus information to be recognized.

Considering the various complications of LDTs, two key decisions were taken to attempt to minimize uncontrolled complexity. Firstly, to avoid the interaction of a different language as L1 with the language of the experiment, it was decided to conduct this experiment on English L1 speakers. Secondly, it was decided to use pseudowords, that is, strings of letters that conform to legal English orthographic and phonological constraints,[5] for example, *nirk*. By using pseudoword primes, it is thus possible to restrict the prime-target relationship to form-based variables only and avoid semantic priming effects. Of course, even English L1 speakers might encounter priming from another language they know, but as they will perceive the experiment to engage with their L1, it should minimize the effect to background noise. Secondly, it is only the masked primes that are pseudowords, whereas the targets are the necessary mixture of real words and pseudowords. As a result, this experiment should come across as fully about the participants' L1.

Participants

Data was collected from 24 English L1 speakers, 20 from the US, 3 from the UK, and

5 The term "pseudoword" is used rather than "nonword" as the latter can refer to an illegal letter string, such as *mgfa*.

1 from Canada (average age 24 years and 8 months, 15 females and 9 males). The participants were all university students in an exchange program at a private university in Japan. A self-report questionnaire indicated a low-intermediate knowledge of Japanese. The participants had normal (or corrected-to-normal) vision and no history of language impairment. The questionnaire also asked if the participants were left- or right-handed. The participants provided informed consent and received a raffle ticket to win a pair of headphones in exchange for their participation. Data was collected over several weeks of university term time. Although 24 participants may appear to be a small sample, Jiang (2012) notes that many reaction time studies are done with similar sample sizes as increasing the number of participants has little effect on the RT data.

Compiling the Stimuli Items for the LDT

In an LDT, two equally sized sets of stimuli are needed for the target items: a set of real words and a set of pseudowords. The inclusion of pseudowords is a necessary part of the design because it prevents the participants from simply pressing the “Yes” button to every single target and developing a bias of responses.

For the set of primes, pseudowords were used to avoid any semantic priming effects. Furthermore, in a task aiming to establish the impact of a feature on how quickly words are recognized, there need to be subsets of words with and without the feature. In this experiment, the set of real word stimuli needed to consist of words used to test the effects of both alliteration and assonance and a control set that had neither feature. The following section outlines how the real words were chosen for the LDT and how the pseudowords were generated for the primes and the “No” response targets.

Real Word Stimuli (Targets)

As adjective-noun collocations and noun-noun compounds were the most commonly used stimuli in the work of Boers et al. (Table 1), it was decided that nouns would also be used as the real word stimuli in the LDT. In an attempt to control for possible frequency, familiarity and length effects, a pool of words was assembled, which comprised 290 monosyllabic nouns of 4–5 letters in length, from the first two frequency sub-lists of the British National Corpus/Corpus of Contemporary American English (BNC/COCA) corpus (Nation, n.d.). Although it is unlikely that any corpus matches the linguistic experience of any particular speaker exactly, this approach was deemed valid in that corpora are representative of the types of input the English L1 participants are likely to have encountered.

The items were then cross-referenced with the Medical Research Council (MRC) database (Coltheart, 1981) for concreteness and imageability ratings. This database contains up to 26 linguistic and psycholinguistic attributes for over 150,000 words collated from previously published sources. The words in the database have subjective ratings on an integer scale from 100 to 700. There were 24 nouns in the initial pool that had no ratings in the MRC database, so these were eliminated from the pool (e.g., *bike*, *dish*, *farm*). The mean and standard deviations were calculated for concreteness and imageability of the remaining 266 nouns. This dataset was then trimmed of any nouns that had ratings of more than ±2 SD from the mean. This resulted in the removal of 11 more nouns (e.g., *bunch, place, thing*), with the aim of producing a range of concreteness and imageability as narrow as possible. The remaining group of 255 potential stimuli had mean concreteness ratings of 576.9 (SD 36.7) and mean imageability ratings of 572.2 (SD 36.1). In comparison, the mean concreteness and imageability ratings in the MRC database are 438 and 450, respectively, suggesting the target stimuli for the LTD were slightly more concrete/imageable than the database average.

The pseudowords in an LDT play two different roles: as primes and as targets for the "No" responses.

Pseudoword Stimuli (Primes)

The nature of an LDT, in which participants have to make rapid decisions as to whether the visual letter string is a real word, means that the characteristics of the pseudowords become an essential part of the experimental design. Not all pseudowords are the same. There is evidence that the types of pseudowords used in a priming experiment have a strong effect on reaction time performance. Keuleers and Brysbaert (2010) state that "the more dissimilar the nonwords are to the words [in the task], the faster are the lexical decision times and the smaller is the impact of word features such as word frequency, age of acquisition, and spelling-sound consistency" (p. 627). Clearly, features such as frequency and age of acquisition apply only to real words, though the form of the pseudoword can affect the speed of all decision processes in an LDT. In Gibbs and Van Orden (1998), for example, mean reaction times to reject nonwords were the shortest when the nonwords were illegal letter strings (e.g., *ldfa*: 496 ms), longer when the stimuli were legal letter strings (e.g., *dilt*: 558 ms), and longer still when the nonwords were pseudohomophones sounding like real words (e.g., *durt*: 698 ms).

To construct matching pseudowords for the 255 nouns, the Wuggy pseudoword generator was used (Keuleers & Brysbaert, 2010) to create monosyllabic stimuli of equivalent length (4–5 letters) with sound-spelling consistency. The Wuggy algorithm produced 10 candidate pseudowords for each real word noun. As words that are orthographically similar to many other words are recognized faster (Yarkoni et al., 2008), neighborhood size and density plays an important role when developing stimuli lists. The Wuggy algorithm uses the Levenshtein edit distance of orthographic similarity (OLD20), where similarity includes neighbors generated by insertion, deletion, substitution, and transposition of letters to produce perceptually similar words. By referring to the OLD20 values in the Wuggy output, it was possible to select the pseudowords that remained as close as possible to real words, that is, with OLD20 values as close to 1.0 as possible. An OLD20 value of 1.8 was arbitrarily set as a cut-off point for pseudoword candidates.

The ideal pseudoword candidate would not cause any inadvertent triggering of semantic or associative representations in a participant's mental lexicon. This proved challenging to control for as a considerable proportion of the 2,550 pseudowords could be construed as a proper name, brand name, acronym, pseudohomophone, or informal, archaic, or variant spelling of a common noun (examples to follow). Moreover, closer scrutiny of the Wuggy output unearthed several words attested in online dictionaries (such as *blog*, *mage*, and *bling*).

The pseudowords were cross-referenced with the COCA database (http://www.americancorpus.org). Any items that appeared as proper names with more than 10 instances in the corpus were discounted (e.g., *crowe*, *imes*, *gide*). Similarly, any items that occurred more than 10 times as acronyms (*pacs*, *facs*, *mact*) were also deleted. Pseudowords were also discounted if they appeared in the corpus as brand names (such as *fage*), abbreviations or variant spellings (e.g., *hols*, *dept*, *lite*, *nite*), or slang (*crip*, *shart*).

The Google search engine was then used to check if the remaining pseudowords elicited any dictionary references, especially in Merriam Webster, as participants were most likely to be from a North American background given the educational context; this led to two more deletions (*thang*, *fleed*). Finally, all pseudohomophones were deleted (e.g., *bocs*, *coaks*, *ceal*, *bild*, *tode*, *wead*).

The experimental design meant that for each of the 255 noun targets, there needed to be one pseudoword prime in each condition: alliteration, assonance, and no phonological pattern. However, after removing all pseudoword candidates with OLD20 values higher than 1.8, pseudohomophones, and those attested in corpora or dictionaries, some nouns did not have equivalent pseudowords in all three conditions; hence, these nouns were removed as test items. Their pseudoword equivalents that met the criteria were retained to use as filler items (see below).

This process of elimination resulted in a final pool of 102 target nouns, each with one alliterating pseudoword prime, one assonating pseudoword prime, and one pseudoword prime with no phonological/orthographic overlap, an example of which can be seen in Table 3.

Table 3
Example of a Stimulus Item Across Three Conditions

Prime	Target	Condition
mouch	*MILK*	Alliteration
skift	*MILK*	Assonance
drate	*MILK*	Control

One limitation of the Wuggy algorithm is that it does not generate the pronunciation for the pseudowords. To check that the prime-target pairs did alliterate and assonate, three English L1 speakers were asked to read aloud the list of items to check the researcher's intuitions of phonological similarity. No discrepancies were found in this regard.

Pseudoword Stimuli (Targets)

To avoid response bias in the LDT, there had to be an equal number of "Yes" and "No" responses. Without such filler items, participants would only need to press the "Yes" key to produce a correct response. Furthermore, such filler items help disguise the critical stimuli, so it is less likely a participant will notice and consequently develop a processing strategy while performing the task. A total of 102 pseudoword targets, which would hopefully elicit a "No" response on the decision task, were selected from the remainder of the Wuggy output. For the 102 pseudoword primes, the mean OLD20 value was 1.34 (min. 1.0, max. 1.7, SD 0.22) and the mean value for the 102 pseudoword fillers was also 1.34 (min. 1.0, max. 1.75, SD 0.24).

In sum, there were 102 noun targets, designed to elicit a "Yes" response in the LDT. The following variables were controlled for frequency, familiarity, word length, neighborhood density, concreteness, and imageability. Each noun target had three pseudoword primes, one in each condition. There were an additional 102 pseudoword targets designed to elicit a "No" response. The primes for these filler items were chosen from the remainder of the Wuggy output. The pseudowords complied with the phonotactic constraints of English and were controlled for in terms of length and orthographic similarity.

Setting up the LDT Program

The LDT was done using the free stimulus presentation software DMDX[6] (Forster & Forster, 2003). An item file was created in a Rich Text Format, which basically instructed the software on what items to present and how to present them. This allowed the 102 nouns plus 102 pseudowords to be combined into 204 trials, which were then divided into 17 blocks. Each block contained two alliterating prime-target pairs, two assonating prime-target pairs, two prime-target pairs with no form overlap, and six pseudoword filler items. Items were scrambled within each block and blocks were also scrambled for each participant. This pseudo-randomization avoided long successions of words or pseudowords appearing by chance. It also ensured extraneous serial effects (such as practice or fatigue) were more evenly distributed across conditions, and no two subjects were likely to receive the same sequence of items. Moreover, it avoided the introduction of systematic errors of measurement arising, when difficult items (on which an error is likely) can affect the reaction time for following items. The three conditions were counterbalanced across three presentation lists in a Latin square design, such that in each list, one third of the stimuli (34 items) appeared in each condition. This allowed for direct comparisons across conditions and avoided list effects. Participants were randomly assigned to one of these lists and responded to an equal number of trials in all three conditions, but never responded to the same target more than once. Two versions of the item file were created to allow for different keys being allocated to "Yes" and "No" responses, one for right-hand dominant participants and one for left-hand dominant participants. The experiment was run on a PC with Microsoft Windows XP (1920×1080 resolution, 60 Hz refresh rate). TimeDX, a component of the DMDX software suite, was used to verify that the hardware features of the computer were satisfactory.

Procedure

Participants were given written instructions (adapted from those used in Ellis et al., 2008) in which they were told that they were participating in a vocabulary experiment and would see a letter string on the computer screen; they were then required to press the key labeled "Yes" if they thought it was an English word and the key labeled "No" if they thought otherwise. Participants were asked to respond as quickly and accurately as possible. Participants were informed neither of the presence of the masked stimuli nor the three critical conditions.

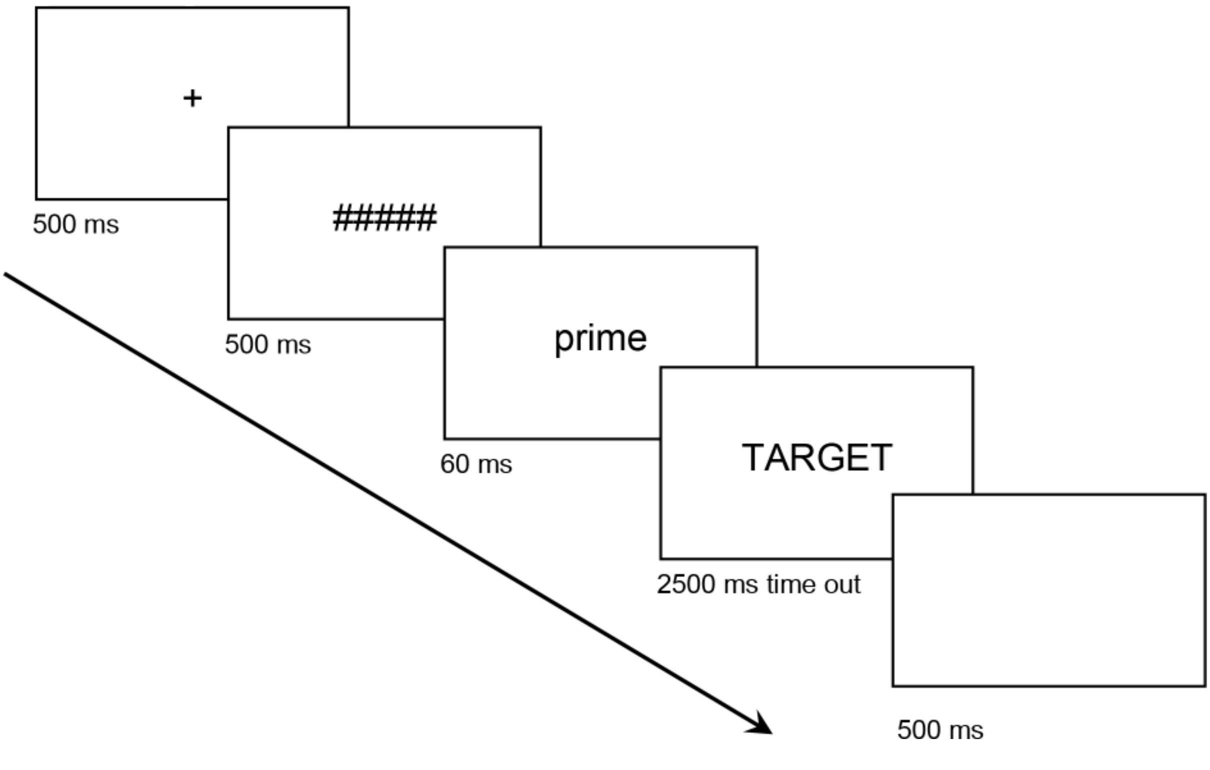
The participants were tested individually and sat approximately 50 cm in front of the

6 Available at http://www.u.arizona.edu/~kforster/dmdx/download.htm

eye-level computer screen, wearing noise-cancelling headphones in a quiet and dimly lit room. The index finger of their dominant hand was assigned to the "Yes" key and their other index finger to the "No" key. The experiment began with 10 practice trials, followed by the 204 experimental trials. There was a programmed break in the middle of the experiment for participants to rest, and the experiment resumed when participants pressed the space key. Trials were randomized for each participant, and the experimental phase took approximately 17 minutes. The stimuli were presented in the center of the computer screen in black Arial font, size 14, on a light gray background using the DMDX software. An explanatory schema of a trial can be seen in Figure 2.

Each trial began with the presentation of a fixation point (+) for 500 ms to direct the attention of the participant to the appropriate location on the screen. A forward mask (a meaningless string of five hash marks) appeared for 500 ms centered at the same location, followed by the presentation of the prime for 60 ms, an amount of time deemed safe for the item to be processed but without the participant becoming aware of it. Then, the target appeared on the screen, which served as a backward mask for the prime. The target remained on the screen until the participant's judgment or a 2500 ms timeout, followed by a 500 ms inter-trial interval (a blank screen). RTs were measured from target onset until the participants' response on the keyboard. After the experiment was completed, participants were asked to complete a short exit questionnaire to self-report on how well they followed task instructions and whether they were aware of the primes.

Figure 2
Sample Trial Schema



Results

Data Analysis

The raw data from the response latencies and associated accuracy were saved as a data file and entered into the *Analyze* program, another component of DMDX, for which an input specification file was written. This text file instructs the program how to treat the raw data and which item goes to what condition. The script parameters also specify how outliers and incorrect responses are to be processed, as described below.

Outliers, defined as responses less than 200 ms or more than 3 standard deviations above the participant's mean, were replaced by the mean value for that participant. This appears to be standard practice in an LDT (McDonough & Trofimovich, 2009). It eliminates the need to perform a log transformation on the subsequent dataset and ensures a normal distribution, thereby safeguarding from the unwanted effects of outlier response latencies.

The script parameters instructed the program to exclude incorrect responses from the analysis. They also instructed the program to identify and automatically remove items that generated error rates of 40% or more, together with the counterpart items in the other two conditions so that test materials remained matched across conditions. The file also instructed the program to reject data from participants with an error rate of 20% or over.

Results

Inspection of the completed data showed that no items generated error rates of 40% or more, so no items were eliminated on these grounds. No participants had error rates of 20% or over. The mean error rate per participant was 3.79% (SD = 2.40), resulting in the removal of 186 incorrect responses. A total of 55 responses qualified as outliers and were replaced by the mean value for that participant (the mean number of outliers per participant was 2.29, SD = 1.23). Furthermore, in the exit questionnaire, none of the participants reported seeing the primes, and all the participants indicated that they had carried out the task exactly as instructed.

A paired-samples t-test was conducted to compare the reaction times in the real word and pseudoword conditions. There was a significant difference in the scores for real word (M = 566.14 ms, SD = 33.20) and nonword (M = 671.8 ms, SD = 76.32) conditions; $t(23) = 6.7$, $p < 0.001$ (two-tailed). These results suggest that the experiment design and choice of pseudoword matches were sound in that words and pseudowords elicited the expected responses, which is in line with Gibbs and Van Orden (1998).

To recap, the research question is whether phonologically patterned prime-target items will result in faster reaction times than prime-targets in the baseline condition with no phonological overlap. Descriptive results from the LDT can be seen in Table 4.

Table 4
Descriptive Results

	Mean Reaction Time (RT) [1]	SD
Alliteration	571.01 ms	34.45
Assonance	585.46 ms	39.07
No Phonological Pattern	575.05 ms	39.92

[1] in milliseconds (ms)

Table 4 shows that alliterating prime-target items produced the fastest mean reaction times and assonating prime-target items the slowest mean reaction times. A repeated-measures ANOVA was used to analyses the results with RTs as the dependent variable and alliteration, assonance, or no pattern as the categorical independent variable. Preliminary examination of the data satisfied the assumptions of homoscedasticity and normality of distribution, and Mauchly's test of sphericity showed that the assumption of sphericity was not violated (χ^2 (2) = 2.93, p = 0.86). There was a significant main effect of phonological pattern on the reaction times ($F_{2, 46}$ = 7.04, p = 0.002, η_p^2 = 0.23, power = 0.91).

Pairwise post hoc comparisons using t-tests and Bonferroni corrected levels of significance showed that participants' RTs were significantly faster in the alliteration condition compared to the assonating condition (mean difference = 14.45, t = 3.69, p = 0.004) with a medium effect size (Cohen's d = 0.76), according to Plonsky and Oswald's (2014) interpretation. However, there was no statistically significant difference between reaction times in the alliteration condition and reaction times in the baseline condition (mean difference = −4.03, t = −1.06, p = 0.90, d = 0.22), nor was there a statistically significant difference between the assonance reaction times and the no phonological pattern reaction times (mean difference = 10.41, t = 2.49, p = 0.06, d = 0.51).

Thus, to answer the research question, phonologically patterned prime-target items did not lead to faster reaction times than prime-targets in the baseline condition with no phonological overlap.

Discussion

The experiment sought to test the hypothesis raised by Boers et al. that MWUs displaying alliteration and assonance have a processing advantage compared to equivalent MWUs with no such phonological or orthographic overlap. Previous priming experiments have found reliable evidence that rhyme produces a priming effect, though the evidence is inconclusive for primes and targets that alliterate, and there is scant data on assonance alone. The results reported here do not contribute to evidence of a facilitatory priming effect with a dataset that used masked pseudoword primes and noun targets. The absence of a statistically significant result could be a consequence of a methodological shortcoming or it could be evidence of the null hypothesis, that is, alliteration and assonance (as operationalized in this experiment) do not facilitate lexical processing.

The experiment used a standard LDT with prime-target items comprising pseudoword primes and monosyllabic high-frequency nouns in three conditions. No individual items generated large error rates, and the participants' mean error rate (5.6%) is in line with other published masked LTDs (e.g., a 6% mean error rate in Perea et al., 2015). The number of participants (N = 24) was deemed appropriate (Jiang, 2012) as was the number of items.[7] Thus, the 24 participants responded to 34 prime-target items in each of the three conditions, generating 816 observations per condition. However, in light of the fact that any mnemonic advantage for alliteration and assonance has tended to be quite modest in the experiments in Table 1, this number of observations per condition may have been insufficient to properly investigate a small effect. Indeed, for a repeated-measures reaction time study, Brysbaert and Stevens (2018) recommend a minimum of 1,600 observations per condition. This suggests that future experiments with sufficient power to detect a small effect would require approximately twice the number of items per condition or twice the number of participants. Recruiting a larger body of participants could perhaps be done more easily via a browser-based application like the *Gorilla Experiment Builder* (Anwyl-Irvine et al. 2020), though there are potential drawbacks to Internet-based reaction time data gathering (see Woods et al., 2015). To conclude, one possible explanation for the lack of a statistically significant finding is that although alliteration and assonance do facilitate lexical processing, the experiment was underpowered and thus was unable to capture the effect.

Another possible reason for not finding a priming effect for phonological patterns relates to the thorny issue of what counts as a "similar sound." In this experiment, alliteration and assonance were operationalized as sounds that match at the phoneme level. This may not be a safe assumption. At this level, allophonic differences created by the phonological environment are not considered. Thus, for example, the /ɪ/ in *skift* and *milk* would be considered the same, even though the latter vowel would be darker before the velarized /l/. Future studies could investigate whether similar phonemes can also alliterate or assonate by virtue of sharing sub-phonemic features.

In sum, the data from the LDT study do not support strong claims that the mnemonic effect of alliteration and assonance is due to a form-based priming effect. This suggests that teachers should draw their students' attention to phonological patterns to aid L2 vocabulary acquisition.

References

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods, 52*(1), 388–407. https://doi.org/10.3758/s13428-019-01237-x

Baddeley, A. (2001). The concept of episodic memory. *Philosophical Transactions of the Royal Society B, 356*(1413), 1345–1350. https://doi.org/10.1098/rstb.2001.0957

Boers, F., Eyckmans, J., & Lindstromberg, S. (2014). The effect of a discrimination task on L2 learners' recall of collocations and compounds. *International Journal of Applied*

7 Jiang (2012) states "It is desirable to have ten to twenty items in a condition whenever it is possible" (p. 48). This experiment had 34 items in each condition.

Linguistics, 24(3), 357–369. https://doi.org/10.1111/ijal.12033

Boers, F., Lindstromberg, S., & Eyckmans, J. (2012). Are alliterative word combinations comparatively easy to remember for adult learners? *RELC Journal, 43*(1), 127–135. https://doi.org/10.1177/0033688212439997

Boers, F., Lindstromberg, S., & Eyckmans, J. (2014a). Is alliteration mnemonic without awareness-raising? *Language Awareness, 23*(4), 291–303. https://doi.org/10.1080/09658416.2013.774008

Boers, F., Lindstromberg, S., & Eyckmans, J. (2014b). When does assonance make L2 lexical phrases memorable? *The European Journal of Applied Linguistics and TEFL, 3*(1), 93–107.

Boers, F., Lindstromberg, S., & Eyckmans, J. (2014c). Some explanations for the slow acquisition of L2 collocations. *Vigo International Journal of Applied Linguistics, 11*(1), 41–62.

Boers, F., Lindstromberg, S., & Webb, S. (2014). Further evidence of the comparative memorability of alliterative expressions in second language learning. *RELC Journal, 45*(1), 85–99. https://doi.org/10.1177/0033688214522714

Brysbaert, M. (2003). Bilingual visual word recognition: Evidence from masked phonological priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: The state-of-the-art* (pp. 323–343). Psychology Press.

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition, 1*(1), 1–20. https://doi.org/10.5334/joc.10

Brysbaert, M., Van Dyck, G., & Van de Poel, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 137–148. https://doi.org/10.1037/0096-1523.25.1.137

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

Crystal, D. (1997). *The Cambridge encyclopedia of language* (2nd ed.). Cambridge University Press.

Dumay, N., Benraïss, A., Barriol, B., Colin, C., Radeau, M., & Besson, M. (2001). Behavioral and electrophysiological study of phonological priming between bisyllabic spoken words. *Journal of Cognitive Neuroscience, 13*(1), 121–143. https://doi.org/10.1162/089892901564117

Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly, 42*(3), 375–396. https://doi.org/10.1002/j.1545-7249.2008.tb00137.x

Eyckmans, J., Boers, F., & Lindstromberg, S. (2016). The impact of imposing processing strategies on L2 learners' deliberate study of lexical phrases. *System, 56*, 127–139. https://doi.org/10.1016/j.system.2015.12.001

Eyckmans, J., & Lindstromberg, S. (2017). The power of sound in L2 idiom learning. *Language Teaching Research, 21*(3), 341–361. https://doi.org/10.1177/1362168816655831

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments and Computers, 35*(1), 116–124.

https://doi.org/10.3758/BF03195503

Gibbs, P., & Van Orden, G. C. (1998). Pathway selection's utility for control of word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 24*(4), 1162–1187. https://doi.org/10.1037/0096-1523.24.4.1162

Goldinger, S. D., Luce, P. A., Pisoni, D. B., & Marcario, J. K. (1992). Form-based priming in spoken word recognition: The roles of competition and bias. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(6), 1211–1238. https://doi.org/10.1037/0278-7393.18.6.1211

Jiang, N. (2012). *Conducting reaction time research in second language studies*. Routledge.

Jusczyk, P. W., Goodman, M. B., & Baumann, A. (1999). Nine-month-olds' attention to sound similarities in syllables. *Journal of Memory and Language, 40*(1), 62–82. https://doi.org/10.1006/jmla.1998.2605

Keuleers, E., & Brysbaert, M. (2010). Wuggy: A multilingual pseudoword generator. *Behavior Research Methods, 42*(3), 627–633. https://doi.org/10.3758/BRM.42.3.627

Lindstromberg, S., & Boers, F. (2008a). The mnemonic effect of noticing alliteration in lexical chunks. *Applied Linguistics, 29*(2), 200–222. https://doi.org/10.1093/applin/amn007

Lindstromberg, S., & Boers, F. (2008b). Phonemic repetition and the learning of lexical chunks: The power of assonance. *System, 36*(3), 423–436. https://doi.org/10.1016/j.system.2008.01.002

Lindstromberg, S., & Eyckmans, J. (2014). How big is the positive effect of assonance on the recall of L2 collocations? *International Journal of Applied Linguistics, 165*(1), 19–45. https://doi.org/10.1075/itl.165.1.02lin

Lindstromberg, S., & Eyckmans, J. (2017). The particular need for replication in the quantitative study of SLA: A case study of the mnemonic effect of assonance in collocations. *Journal of the European Second Language Association, 1*(1), 126–136. https://doi.org/10.22599/jesla.26

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review, 88*(5), 375–407. https://doi.org/10.1037/0033-295X.88.5.375

McDonough, K., & Trofimovich, P. (2009). *Using priming methods in second language research*. Routledge.

McQueen, J. M., & Sereno, J. (2005). Cleaving automatic processes from strategic biases in phonological priming. *Memory & Cognition, 33*(7), 1185–1209. https://doi.org/10.3758/BF03193222

Monsell, S., & Hirsh, K. W. (1998). Competitor priming in spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(6), 1495–1520. https://doi.org/10.1037/0278-7393.24.6.1495

Nakayama, M., Sears, C. R., Hino, Y., & Lupker, S. J. (2012). Cross-script phonological priming for Japanese-English bilinguals: Evidence for integrated phonological representations. *Language and Cognitive Processes, 27*(10), 1563–1583. https://doi.org/10.1080/01690965.2011.606669

Nation, I. S. P. (n.d.). The BNC/COCA word family lists. https://www.wgtn.ac.nz/__data/assets/pdf_file/0005/1857641/about-bnc-coca-vocabulary-list.pdf

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.

Nation, I. S. P. (2014). *What do you need to know to learn a foreign language?* Victoria University of Wellington. https://www.wgtn.ac.nz/lals/resources/paul-nations-resources/paul-nations-publications/publications/documents/foreign-language_1125.pdf

Norris, D., McQueen, J. M., & Cutler, A. (2002). Bias effects in facilitatory phonological priming. *Memory & Cognition, 30*(3), 399–411. https://doi.org/10.3758/BF03194940

Perea, M., Vergara-Martinez, M., & Gomez, P. (2015). Resolving the locus of case alternation effects in visual word recognition: Evidence from masked priming. *Cognition, 142*, 39–43. https://doi.org/10.1016/j.cognition.2015.05.007

Peters, A. M. (1983). *The units of language acquisition* (Vol. 1). CUP Archive.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning, 64*(4), 878–912. https://doi.org/10.1111/lang.12079

Radeau, M., Morais, J., & Dewier, A. (1989). Phonological priming in spoken word recognition: Task effects. *Memory & Cognition, 17*(5), 525–535. https://doi.org/10.3758/BF03197074

Rastle, K., & Brysbaert, M. (2006). Masked phonological priming effects in English: Are they real? Do they matter? *Cognitive Psychology, 53*(2), 97–145. https://doi.org/10.1016/j.cogpsych.2006.01.002

Rubin, D. C. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads and counting-out rhymes*. Oxford University Press.

Shallice, T., & Cooper, R. P. (2011). *The organisation of mind*. Oxford University Press.

Slowiaczek, L. M., & Hamburger, M. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(6), 1239–1250. https://doi.org/10.1037/0278-7393.18.6.1239

Slowiaczek, L. M., Nusbaum, H. C., & Pisoni, D. B. (1987). Phonological priming in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*(1), 64–75. https://doi.org/10.1037/0278-7393.13.1.64

Slowiaczek, L. M., & Pisoni, D. B. (1986). Effects of phonological similarity on priming in auditory lexical decision. *Memory & Cognition, 14*(3), 230–237. https://doi.org/10.3758/BF03197698

Szudarski, P. (2017). Learning and teaching L2 collocations: Insights from research. *TESL Canada Journal, 34*(3), 205–216. https://doi.org/10.18806/tesl.v34i3.1280

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

Woods, A. T., Velasco, C., Levitan, C. A., Wan, X., & Spence, C. (2015). Conducting perception research over the internet: A tutorial review. *PeerJ*, *3*, e1058. https://doi.org/10.7717/peerj.1058

Yap, M. J., & Balota, D. A. (2015). Visual word recognition. In A. Pollatsek & R. Treiman (Eds.), *The Oxford handbook of reading* (pp. 26–43). Oxford University Press.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's *N*: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15*(5), 971–979. https://doi.org/10.3758/PBR.15.5.971

Zwitserlood, P. (1996). Form priming. *Language and Cognitive Processes, 11*(6), 589–596. https://doi.org/10.1080/016909696386980