Check for updates

Costs and Benefits of Native Language Similarity for Non-native Word Learning

Viorica Marian[1]*, James Bartolotti[2], Aimee van den Berg[1] and Sayuri Hayakawa[1]

[1] Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL, United States, [2] Life Span Institute, University of Kansas, Lawrence, KS, United States

The present study examined the costs and benefits of native language similarity for non-native vocabulary learning. Because learning a second language (L2) is difficult, many learners start with easy words that look like their native language (L1) to jumpstart their vocabulary. However, this approach may not be the most effective strategy in the long-term, compared to introducing difficult L2 vocabulary early on. We examined how L1 orthographic typicality affects pattern learning of novel vocabulary by teaching English monolinguals either Englishlike or Non-Englishlike pseudowords that contained repeated orthographic patterns. We found that overall, the first words that individuals learned during initial acquisition influenced which words they acquired later. Specifically, learning a new word in one session made it easier to acquire an orthographically similar word in the next session. Similarity among non-native words interacted with native language similarity, so that words that looked more like English were easier to learn at first, but they were less effective at influencing later word learning. This demonstrates that although native language similarity has a beneficial effect early on, it may reduce learners' ability to benefit from non-native word patterns during continued acquisition. This surprising finding demonstrates that making learning easier may not be the most effective long-term strategy. Learning difficult vocabulary teaches the learner what makes non-native words unique, and this general wordform knowledge may be more valuable than the words themselves. We conclude that native language similarity modulates new vocabulary acquisition and that difficulties during learning are not always to be avoided, as additional effort early on can pay later dividends.

Keywords: language learning, cross-language similarity, second language, language acquisition, vocabulary learning

INTRODUCTION

Children often excel at learning new languages—consider international adoptees who rapidly acquire their "second first language" (Roberts et al., 2005)—whereas for adults, learning a second language (L2) has traditionally been thought to be a more difficult task (Liskin-Gasparro, 1982). There is now substantial evidence that, for children and adults alike, the ability to successfully learn a second language can be moderated by complex interactions between contextual, sociocultural, cognitive, and affective variables (see Dixon et al., 2012 and Ortega, 2013 for reviews), as well as characteristics of the first (L1) and second (L2) languages. Oftentimes, learners can take advantage of similarities between the L1 and L2 by relying on existing skills and knowledge to learn the

new language (i.e., cross-linguistic transfer or cross-linguistic influence; Ringbom, 2007; Jarvis and Pavlenko, 2008). Other times, transfer from the L1 can inhibit learning, such as when L1 knowledge is inappropriately applied (Laufer, 1988; Eckman, 2004) or when it interferes with the acquisition of L2-specific representations (Flege, 1987; Goldrick et al., 2014). Seeing language acquisition as an incremental process (that successively builds on previously-learned information), the costs and benefits of cross-linguistic influences at early stages of learning could have cascading consequences for later acquisition. In the current study, we examine the developmental trajectory of cross-linguistic influences on novel word learning and the role of orthographic similarity to previously-learned native and non-native words.

Effects of Cross-Linguistic Transfer on New Word Learning

Native language similarity has long been known to be a powerful resource for language learning. In many cases, the ease and efficiency of language acquisition can be modulated by existing language knowledge and the actual (Ringbom, 2007) or perceived (Kellerman, 1978; Odlin, 1989) formal similarities between languages. Similarities between languages and cross-linguistic transfer can be found at multiple levels of representation, such as phonology (e.g., Melby-Lervåg and Lervåg, 2011; Wrembel, 2011), orthography (e.g., De Groot and Keijzer, 2000; Ellis, 2008), and morphology (e.g., Hancin-Bhatt and Nagy, 1994; Ecke, 2015). During vocabulary learning, cognates, which overlap across languages in both orthographic form and meaning, are often more readily acquired than non-cognates (Lotto and de Groot, 1998; De Groot and Keijzer, 2000). In addition to allowing learners to draw from existing knowledge, L1-L2 similarity may facilitate integration of novel wordforms into the existing lexico-semantic network (Shirai, 1992; MacWhinney, 1997; Comesaña et al., 2009; De Groot, 2011), which can result in more robust encoding during early stages of acquisition (Ellis and Beaton, 1993), as well as more fluid retrieval at higher levels of proficiency (e.g., Comesaña et al., 2012).

SLA studies conducted in more naturalistic contexts (e.g., classrooms) have also found advantages for learning cognates (e.g., Cunningham and Graham, 2000; Tonzar et al., 2009; Vidal, 2011; Otwinowska and Szewczyk, 2019; Puimège and Peters, 2019), but have produced more mixed results (e.g., Rogers et al., 2015; Otwinowska et al., 2020; see Otwinowska, 2015 for review). Some evidence suggests that cognate facilitation may be contingent on formal training in cognate recognition (Tréville, 1996; Dressler et al., 2011), suggesting that learners may not always be aware of formal similarities. Indeed, using contrastive analysis to highlight similarities and differences between the L1 and L2 can be highly effective (Laufer and Girsai, 2008; Lin, 2015; Helms-Park and Perhan, 2016), and students tend to respond positively to this type of language instruction (Brooks-Lewis, 2009). Additionally, the trade-off between ecological validity and control over stimulus characteristics (e.g., word frequency, orthographic overlap) may also contribute to the more tenuous cognate effects found in SLA studies compared to laboratory

experiments (but see Otwinowska and Szewczyk, 2019 for an exception and Otwinowska, 2015 for discussion).

Psycholinguistic studies using carefully controlled real or artificial word stimuli have revealed that the cognate advantage increases with the degree of orthographic overlap (De Groot, 2011; Comesaña et al., 2015), and that even without semantic overlap, vocabulary acquisition can be facilitated for novel words that are orthographically (De Groot, 2011; Bartolotti and Marian, 2014, 2017; Marecka et al., 2021) or phonologically (Ellis and Beaton, 1993; Service and Craik, 1993; Roodenrys and Hinton, 2002; Storkel and Maekawa, 2005; Storkel et al., 2006) similar to L1 words. For instance, Meade et al. (2018) observed that pseudowords with a higher number of L1 orthographic neighbors were produced more accurately than low-density words. Like cognates, the advantage for words that resemble L1 wordforms could result from more effective use of, and integration with, existing lexical and semantic knowledge. This can be done explicitly through association-based strategies (e.g., mneumonic methods; Atkinson and Raugh, 1975; Meara, 1980; Paivio and Desrochers, 1981; see Hulstijn, 1997 and Nation, 1982 for reviews) or implicitly through the co-activation of orthographically or phonologically similar L1 words and associated meanings (Holcomb et al., 2002; Van Hell and Tanner, 2012).

Words with familiar orthographic or phonological features may additionally benefit word learning by allowing learners to exploit knowledge of sublexical regularities in the L1. For instance, Bartolotti and Marian (2014) found that vocabulary acquisition is facilitated when pseudowords are designed to reuse native language letter patterns (i.e., higher bigram probabilities; see also Bartolotti and Marian, 2017). Phonological overlap between the L2 and L1 can additionally facilitate learning by increasing pronounceability (Ellis and Beaton, 1993; Service and Craik, 1993), which could enable learners to rely on phonological knowledge stored in long-term memory (Cheung, 1996; Storkel and Maekawa, 2005; De Groot, 2011) and the mental rehearsal of novel phonological forms (Papagno et al., 1991; Ellis and Sinclair, 1996).

Despite potential advantages, lexical and sublexical similarities between languages can also introduce costs when they are over-applied or block acquisition of new features during learning. For example, a German learner of English may say, "I need a loffel for my soup," under the mistaken belief that the German word *Löffel* (meaning spoon) is an English cognate (Eckman, 2004). This type of confusion can be especially likely when a novel word overlaps with a known word in some respects, but not others (i.e., "deceptive transparency;" Laufer, 1988), such as when an L2 word overlaps in form but not meaning with an L1 word (i.e., false friends; e.g., the German word *Rat*, which means "advice"). A different problem occurs when similarity to the L1 interferes with acquisition of L2-specific features or regularities, as seen with spoken accents. Sounds in the L2 that are similar to an existing L1 sound are actually *more* difficult to pronounce accurately than completely new sounds (Flege, 1987). Even speakers who have mastered L2 phonology still pronounce cognate words with more of an accent than non-cognates, due to cognates' high L1 similarity (Amengual, 2012,

2016; Goldrick et al., 2014; consistent with Anderson's (1983) *transfer to somewhere* principle).

Kaushanskaya and Marian (2009) found that, compared to when phonologically atypical pseudowords were presented alone, learning was impaired when words were presented bimodally with *typical* L1 orthographic forms. Such effects could be explained by the increased activation of L1 representations in response to familiar wordforms, which could compete with the more recently acquired L2 representation. For instance, psycholinguistic studies have demonstrated that visual word recognition can be inhibited by orthographically-related primes (both within and across languages; e.g., Bijeljac-Babic et al., 1997) that can compete for selection. Cross-language lexical activation can be reduced, however, by priming different script bilinguals (e.g., Hindi-English) with a particular writing system (Dubey et al., 2018), as well as by priming same script bilinguals with language-specific sublexical cues (e.g., bigrams) that are uncommon or orthotactically illegal in one language but not the other (Casaponsa et al., 2020). Similarly, language discrimination is facilitated by orthographic markers that signal language membership (Casaponsa et al., 2014; Oganian et al., 2015). These patterns of facilitation and interference indicate a high degree of cross-linguistic interactivity within the language system, which can play a significant role in vocabulary acquisition and be amplified by similarities at the lexical and sublexical level of processing.

Effects of Within-Language Transfer on New Word Learning

Similarity to the L1 can yield significant benefits during early stages of word learning by encouraging cross-linguistic transfer. The recognition and use of L2-specific patterns, however, is key to long-term success in developing L2 vocabulary. Adults who had completed 1 year of university-level Spanish courses were able to learn new words with a large number of Spanish neighbors (i.e., words that differed from many Spanish words by only a single phoneme) at a higher rate than words with a low number of Spanish neighbors (Stamer and Vitevitch, 2012). This ability to learn words with more L2 neighbors provides evidence that similarities *within* an L2 benefit learning. The application of within-language knowledge for novel word learning can additionally vary as a function of individual differences such as L2 proficiency (e.g., Horst et al., 1998; Zahar et al., 2001; Pulido, 2003; Tekmen and Daloglu, 2006; Ma et al., 2015; Otwinowska and Szewczyk, 2019). For instance, Ma et al. (2015) observed that L2 proficiency was positively associated with learners' ability to learn the meanings of novel pseudowords embedded in sentences. Studies employing incidental learning paradigms have similarly observed that higher L2 proficiency and larger L2 vocabulary sizes facilitate novel vocabulary acquisition during reading (e.g., Horst et al., 1998; Tekmen and Daloglu, 2006). Such findings suggest that as proficiency in the L2 increases, so too does the strength of within-L2 facilitation, creating a positive feedback loop where L2 word learning becomes easier as L2 vocabulary size increases. Proficiency can also modulate the contribution of other domain-general cognitive abilities

(Cheung, 1996; Gathercole and Masoura, 2005). For instance, Cheung (1996) found that greater short-term memory capacity was associated with better L2 vocabulary learning for individuals with low, but not high L2 proficiency (but see Majerus et al., 2008 who found independent effects of STM and L2 phonological knowledge). Bartolotti et al. (2011) observed that inhibitory control and bilingual experience independently predicted how well learners were able to extract statistical regularities of word boundaries in an artificial Morse Code language after listening to another language with conflicting patterns. Bilingual experience facilitated word learning when interference was low, whereas inhibitory control predicted performance when interference was high (see also Wang and Saffran, 2014, who observed a bilingual advantage for detecting regularities in an artificial tonal language). The ability to extract and apply regularities within the L2 can therefore vary depending on both individual differences in cognitive and linguistic abilities, as well as characteristics of the learning task.

Though gains are likely to compound with increased L2 experience, the beneficial effect of within-L2 similarity applies even at the earliest stages of acquisition (McLaughlin et al., 2004; Bartolotti and Marian, 2017). After only 14 h of classroom study, novice L2 learners' neural responses indicated familiarity with words they had seen before, even though behaviorally they only identified words at chance performance (McLaughlin et al., 2004; Osterhout et al., 2006). After only one session of training in an artificial language, learners demonstrate that they have learned letters' relative frequencies in the language, and can use this information to fill gaps in their knowledge of the new language (Bartolotti and Marian, 2017). Other statistical regularities governing word boundaries can be learned from continuous speech after as little as 20 min of exposure (Saffran et al., 1999; Karuza et al., 2013), and this knowledge of word boundaries can directly influence subsequent vocabulary acquisition (Mirman et al., 2008). Together, these findings demonstrate that learners are able to extract L2 regularities based on even brief amounts of exposure, which can then be used to support further learning.

Effects of Cross-Linguistic Influence on Within-Language Transfer

While there has been substantial research investigating the independent effects of between- and within-language transfer on vocabulary acquisition, relatively less is known about their potential interactions—specifically, whether native language orthographic similarity modulates transfer between non-native words during subsequent learning. A significant body of research has shown that the strategies individuals use to process words within the L1 is influenced by their orthographic system (e.g., Hakuta, 1982; MacWhinney and Bates, 1989), and that these same processes may be used to decode words in an L2 (e.g., Koda, 1998; Mori, 1998; Hamada and Koda, 2008). As a result, sensitivity to L2-specific orthotactics can vary as a function of similarity between L1 and L2 orthographic systems. For instance, Koda (1998) observed that ESL learners with L1 Korean (who utilize a syllable-based writing system, *hangul*) were more sensitive to English orthotactic violations (i.e.,

illegal letter sequences) than L1 Chinese speakers (who utilize a morpheme-based logographic writing system). The author conjectures that the Korean speakers' increased sensitivity to L2 intraword structures likely results from their greater need to attend to component letters and sequences during L1 decoding relative to Chinese speakers. This finding suggests that L1 experience can modulate learning of L2-specific regularities, with variable outcomes depending on how well the strategies acquired for the L1 can be applied to the L2 (see also Koda, 1990, 1993 for similar effects of L1 orthography and transfer on L2 reading comprehension strategies and Koda and Zehler, 2008, for review). The present study examines the possibility that effects of cross-linguistic similarity may be observed when languages overlap, not only in their orthographic systems as a whole, but in the orthographic forms of particular words.

Preliminary support for this possibility comes from another study by Koda (1989), who found that Japanese L2 learners whose L1 (Korean or Chinese) overlapped with one type of Japanese script (kanji, a logographic, meaning-based system), but not another (hiragana, a phonetic lettering system), outperformed learners with orthographically dissimilar L1s in learning vocabulary of both scripts. Furthermore, initial L1-similarity advantages for word learning compounded to yield later benefits for more complex tasks, such as reading comprehension. While these findings indicate that L1 similarity for a subset of L2 vocabulary can facilitate the acquisition of other L2-specific wordforms (potentially via transfer of phonological representations that map to both kanji and hiragana), learners with knowledge of logographic characters could have benefited from overlap in both orthographic form and meaning (akin to cognate facilitation). The present study therefore examines whether similar benefits of cross-linguistic influence on within-language transfer can be observed when a subset of novel words overlap with the L1 in sublexical properties alone.

Given that adult language learners' primary approach when they start learning a new language is typically to identify and reuse perceived similarities to their native language (Ringbom and Jarvis, 2011), it would be consequential to know how increased activation and use of L1 knowledge during early stages of learning affects learners' ability to later rely on regularities within the L2. One possibility is that identifying useful similarities between the L1 and L2 during initial acquisition will enhance the ability to learn and use similarities within the L2. For instance, learning L2 words that share orthographic features with the L1 could establish a stronger base of knowledge to be used as exemplars for subsequently learned words with L2-specific features. In addition to potential differences in the strength of L2 (exemplar) representations, the cognitive processes and strategies engaged while learning words that resemble the L1 could increase the salience and use of L2 regularities.

There are also reasons to expect that transfer between L2 words could instead be facilitated by the initial acquisition of wordforms that are dissimilar to the L1. For instance, the greater challenges associated with learning dissimilar words could serve as a form of desirable difficulty (Bjork and Bjork, 2011), which could elicit higher levels of involvement (Craik and Lockhart, 1972; Laufer and Hulstijn, 2001; Rice and Tokowicz, 2020)

or motivation (Dörnyei and Ushioda, 2009; Dörnyei, 2019), resulting in deeper processing and greater sensitivity to L2-specific patterns. It may also be the case that words with less typical L1 orthography would elicit relatively less activation of L1 representations that could interfere with the identification and use of L2 features (e.g., Amengual, 2012, 2016; Goldrick et al., 2014). If so, we may observe greater within-L2 transfer after learning words with orthographic features that are uncommon in the L1.

In the present study, we introduce the concept of "bridge" words as a means to investigate cross-linguistic influences on transfer between novel vocabulary and the potential utility of bridge words for teaching learners about useful features of non-native words. Bridge words are defined as novel words that contain letter sequences that are common among the non-native vocabulary to facilitate subsequent learning. Acquiring a bridge word (e.g., haner in the current study) may make it easier to learn a similarly spelled "terminus" word (e.g., hajer) to which it is connected because of orthographic feature overlap. Some bridge words use letter sequences that are also common in the L1, which may make them easier to acquire, whereas other bridge words have orthographic forms that are uncommon in the L1. To examine the effect of L1 similarity on bridge words' utility, we designed contrasting sets of pseudowords and taught participants one of two word lists across two sessions. Participants were first taught bridge words comprised of letter sequences (i.e., bigrams) that were either typical (i.e., "Familiar;" e.g., haner, meaning "bride") or atypical of English words (i.e., "Unfamiliar;" e.g., vobaf, meaning "cloud"), followed by an immediate test where they produced the new word when cued with its meaning. Two weeks later, participants returned to learn terminus words that were related to their previously-learned bridge words (e.g., hajer, tobaf), and were again tested immediately. If we observe a general benefit for terminus word acquisition based on bridge word knowledge, it would suggest that learners are able to use orthographic similarities within the non-native vocabulary to facilitate subsequent learning. Critically, if we observe different effects of bridge words in the Familiar and Unfamiliar conditions, it would suggest that native language orthotactic typicality can modulate how knowledge specific to non-native words is used. Native language similarity may improve bridge-to-terminus transfer, by accentuating word-to-word similarity as a learning tool, or it may interfere, by hindering acquisition of non-native patterns.

METHODS

Participants

A power analysis to determine sample size was run with Monte Carlo simulations using the SIMR package in R for use with linear mixed effect models (Green and Macleod, 2016). An effect size for the influence of L1 orthographic bigram typicality on learning was obtained from word learning data in Bartolotti and Marian (2017), providing a fixed effect estimate of 10% on learning accuracy. Population mean and variance were obtained from pilot data. Power estimates were calculated for simulated

sample sizes from 20 to 40. Power >0.8 was obtained with 30 participants, and power >0.9 was obtained with 38 participants.

Sixty-five English-speaking adults initially participated after providing informed consent in accordance with the university's institutional review board, and were randomly assigned to learn Familiar or Unfamiliar word lists. Participants' language profiles were collected using the LEAP-Q (Marian et al., 2007). Non-verbal IQ was assessed using the matrix reasoning subtest of the Wechsler Abbreviated Scale of Intelligence (PsychCorp, 1999). Verbal memory was assessed using the verbal paired associates test of the Wechsler Memory Scale III (Wechsler, 1997).

As the novel vocabulary used in the present study was controlled for orthographic wordform similarity to English (i.e., bigram and biphone probability), but not other languages, only native English-speakers with minimal second language knowledge were included in the final sample. Eligible participants had self-reported second language proficiencies (speaking, listening, and reading composite score) of less than 3 (corresponding to "low" proficiency) on a scale of 0–10 (ranging from "none" to "perfect"). This language knowledge criterion was applied prior to data analysis and excluded all bilinguals and multilinguals ($N = 27$), yielding a final sample size of 38 participants (Familiar group $N = 17$, Unfamiliar group $N = 21$). Participants in the Familiar and Unfamiliar groups did not differ in non-verbal IQ standard scores (Familiar $M = 111.0$, $SE = 0.55$, Unfamiliar $M = 110.5$, $SE = 0.38$, $t(28.1) = 0.17$, n.s.) or verbal memory standard scores (Familiar $M = 13.0$, $SE = 0.15$, Unfamiliar $M = 13.4$, $SE = 0.13$, $t(32.6) = 0.45$, n.s.). All participants were students at Northwestern University who completed the study in a classroom-like setting in exchange for extra credit.

Materials

The Familiar and Unfamiliar word lists each contained 48 five-letter words with alternating consonants and vowels in CVCVC format (Q, Y, and X were not used in either language). Two versions of each word list were created, one per training session. Vocabulary items in the first list were used to examine L1 influences on learning, and were selected by evaluating 10,000 randomly generated non-words for English similarity. Though word lists were presented visually, psycholinguistic evidence suggests that phonological forms of words are co-activated even in response to unimodal orthographic inputs (e.g., Perfetti and Bell, 1991; Ferrand and Grainger, 1993; Grainger and Ferrand, 1994; Brysbaert et al., 1999; Van Wijnendaele and Brysbaert, 2002; Brysbaert and Van Wijnendaele, 2003; Grainger et al., 2006; Braun et al., 2009). English similarity was therefore determined based on both bigram and biphone probabilities. Phonological forms of each novel word were determined using the eSpeak speech synthesizer software, version 1.48.15 for Linux (Duddington, 2012). Pronunciations were obtained as IPA transcriptions using eSpeak's EN-US American English voice, and were translated from IPA to the CPSAMPA format (a modification of XSAMPA) for use with CLEARPOND (Marian et al., 2012). The orthographic and phonological forms of each novel word were used to obtain average bigram and biphone probabilities in English, and English similarity was

defined as a composite metric of z-transformed bigram and biphone probabilities.

To establish high and low English similarity thresholds, an English similarity percentile rank score was defined based on real English words. All five-letter English words in SUBTLEX-US (Brysbaert and New, 2009) with a frequency-per-million of 0.33 or greater were used to create the English similarity score. Each real word's score (i.e., average of z-transformed English bigram and biphone probabilities) was calculated and words were rank-ordered by English similarity. A High English similarity threshold was defined at the 20th percentile score, and 48 of the randomly generated novel words with scores above the threshold were selected for the first Familiar word list. A Low English similarity threshold was defined at the 99th percentile score, and 48 of the novel words with scores below the threshold were selected for the first Unfamiliar word list. Words in both lists were selected with the additional constraint of ensuring a balanced distribution of letters at word onset.

An additional 48 novel words in each condition (Familiar, Unfamiliar) were designed for use in the second session, which examined the effect of similarity to previously-learned words on new word learning. All new "terminus words" in the second session were substitution neighbors of a single item from that condition's "bridge word" list, learned in the first session. New terminus words were selected from a list comprising all non-word single-letter substitution neighbors of entries from the bridge word list (excluding duplicate entries, which were neighbors of multiple words in the bridge word list). In order to assess how well learners are able to utilize non-native patterns to learn other new words, English similarity was calculated for all generated entries and only new terminus words with scores below the Low English similarity threshold were selected for *both* the Familiar and Unfamiliar conditions. In other words, while the Familiar and Unfamiliar bridge word lists differed in English bigram/biphone probability for the first session, terminus words in the second session were equally dissimilar to English, thereby ensuring that effects of condition observed for terminus words could not be attributed to direct transfer from the L1. From this reduced list, 48 terminus words were randomly selected for each condition, with the constraints that each terminus word was a neighbor of a different word from the bridge word list and that the average English bigram/biphone probability did not differ between the second lists in each condition or between the bridge and terminus lists in the Unfamiliar condition (all ps > 0.1; see **Supplementary Tables 1**, **2** for bridge and terminus wordform statistics and stimuli).

All novel words were assigned a different English meaning for use during learning; the Familiar and Unfamiliar conditions both used the same list of 96 English words. To control for effects of individual novel-word—English-word pairings, two variants were created for each condition. The 96 English words were divided into A and B lists that each included equal numbers of concrete (e.g., "tree") and abstract (e.g., "idea") nouns (as determined by measures of imageability, see De Groot, 2006 for a similar approach). The two lists were matched for imageability, age of acquisition, and familiarity (Bristol norms) (Stadthagen-Gonzalez and Davis, 2006), as well as lexical frequency on

the SUBTLEX-US zipf scale (Brysbaert and New, 2009; Van Heuven et al., 2014) (*p*s > 0.05; see **Supplementary Tables 3, 4** for English words in lists A and B, as well as statistical comparisons of word characteristics in the two lists). For half of the participants in each language group, list A meanings were assigned to novel words in the bridge session and list B was used for the terminus session, while the other half of participants received list B meanings in the bridge session and list A meanings in the terminus session. In this way, each participant learned a translation of the same 96 English words (with each English word paired with a single novel word), with imageability, age of acquisition, familiarity, and lexical frequency controlled across the four list types (Familiar-Bridge, Familiar-Terminus, Unfamiliar-Bridge, Unfamiliar-Terminus). Lastly, in order to account for possible differences between groups in similarity between the novel wordforms and the wordforms of their direct English translations (e.g., a cognate effect or near-cognate effect), we confirmed that the number of novel word—English word pairs that had overlapping bigrams (e.g., a novel word "cohuz" paired with the English word "command") did not differ between the Familiar and Unfamiliar Bridge word lists (2 and 0 out of 96, respectively) or between the Familiar and Unfamiliar Terminus word lists (1 out of 96 in both, *p*s > 0.05).

Procedure

Participants learned the novel bridge and terminus word lists they were assigned over the course of two sessions spaced 2 weeks apart. In each session, each participant was given a sheet of paper containing all 48 novel bridge or terminus words and their meanings printed as paired associates (e.g., *haner*—bride). Participants were provided 16 min to silently learn as many words as they could, and were told that they would be tested immediately afterwards. While the use of a more structured task (e.g., timed presentation of individual word pairs) can be beneficial for isolating the mechanisms underlying effects on learning, the present study was designed to be an initial test of the hypothesis that similarity to native language words would modulate transfer of non-native knowledge. The use of carefully controlled word stimuli combined with a self-paced paired-associates task enabled us to assess the overall impact of native language similarity on non-native transfer without imposing constraints on learners' allocation of time to study individual words. This approach additionally allowed us to simultaneously test groups of participants in a classroom-like setting using a format commonly found in foreign language textbooks and study materials (see Prince, 1996; Laufer and Shmueli, 1997; Hermann, 2003; Webb, 2007 for similar approaches). The duration of the study phase was determined based on pilot data and prior studies utilizing similar paradigms (e.g., Pickering, 1982; Prince, 1996; Laufer and Shmueli, 1997; Webb, 2007). Following the study phase, participants were then given 6 min to write the matching novel word translations on a response sheet containing all 48 English meanings. The order of words was fixed across participants but randomized between learning and test. A research assistant later manually transcribed written responses onto a computer, which automatically scored participants' accuracy.
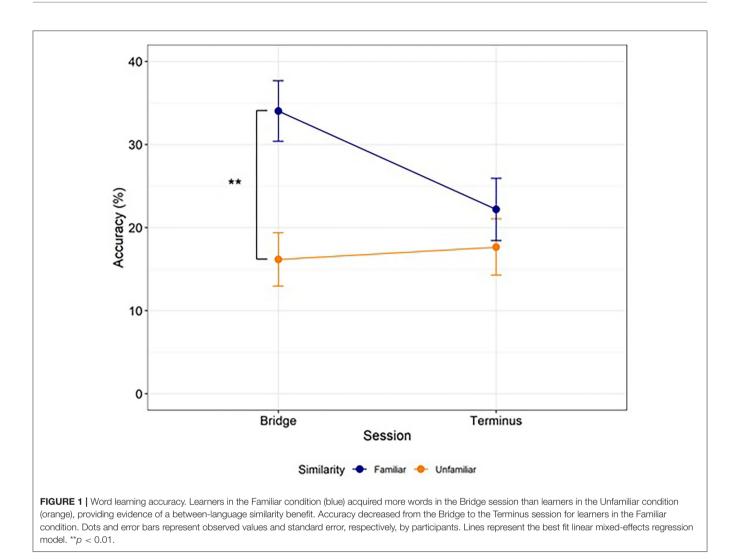
Data Analysis

Response accuracy was calculated taking into account partially correct responses. Each correct letter in the correct position of a response scored 0.2 points, for a maximum score of 1. The effects of native and non-native word similarity on accuracy were analyzed with linear mixed effects-regression, using the lme4 package (Bates et al., 2014) in R (R Core Team, 2016). Models included fixed effects of Similarity Condition (Familiar, Unfamiliar) and Session (Bridge word, Terminus word), plus an interaction term. Imageability, familiarity, age of acquisition, and word frequency of the English translations were added as covariates. The models additionally included random intercepts for participants, word forms, and word meanings, allowing us to control for mean learning performance associated with individual participants and words. Models additionally included a by-participant random slope for Session and by-meaning random slopes for Session and Similarity (i.e., the "maximal" random effects structure[1], Barr et al., 2013), allowing us to control for random variation in the fixed effects associated with individual participants and words. Significance of fixed effect estimates was evaluated using the Satterthwaite approximation for degrees of freedom. Follow-up comparisons on models' predicted marginal means (using Welch *t*-tests) also used the Satterthwaite approximation for degrees of freedom, and the Tukey correction for multiple comparisons.

RESULTS

We found a significant interaction between Similarity and Session [*Estimate* = −13.35, *SE* = 4.93, 95% CI (−23.01, −3.69), *t*(51.65) = −2.71, *p* = 0.009], as well as a main effect of Similarity [*Estimate* = 11.37, *SE* = 4.26, 95% CI (3.02, 19.72), *t*(43.98) = 2.67, *p* = 0.011] and a marginally significant main effect of Session [*Estimate* = −4.44, *SE* = 2.54, 95% CI (−9.42, 0.54), *t*(58.32) = −1.75, *p* = 0.086] (**Figure 1**). Follow-up comparisons on the model's predicted marginal means revealed that accuracy for the Familiar condition in the Bridge session *M* = 34.03, *SE* = 3.65, 95% CI (26.88, 41.19) was higher than for the Unfamiliar condition in the Bridge session *M* = 16.16, *SE* = 3.22, 95% CI (9.85, 22.48), *z* = −3.69, *p* = 0.001, and higher than accuracy for either the Familiar condition *M* = 22.18, *SE* = 3.74, 95% CI (14.85, 29.52), *z* = 3.2, *p* = 0.007 or the Unfamiliar condition in the Terminus session *M* = 17.66, *SE* = 3.38, 95% CI (11.03, 24.29), *z* = 3.26, *p* = 0.006. No other comparisons were significant.

The higher accuracy in the Bridge session for the Familiar condition compared to the Unfamiliar condition demonstrates a substantial benefit of native language similarity during self-directed vocabulary learning. However, the better learning observed for the Familiar condition did not carry through to the subsequent Terminus session, at which point there was no

---

[1]Note that no random slopes for Similarity were included for participants or word forms as each participant and word form was assigned to a single similarity condition (either Familiar or Unfamiliar). No random slope for Session was included for word form as each word form was presented in a single session (either bridge or terminus).

FIGURE 1 | Word learning accuracy. Learners in the Familiar condition (blue) acquired more words in the Bridge session than learners in the Unfamiliar condition (orange), providing evidence of a between-language similarity benefit. Accuracy decreased from the Bridge to the Terminus session for learners in the Familiar condition. Dots and error bars represent observed values and standard error, respectively, by participants. Lines represent the best fit linear mixed-effects regression model. ** $p < 0.01$.

significant difference between word retrieval accuracy in the two groups.

The Terminus session contained entirely new vocabulary for participants to learn; all words were single letter substitution neighbors of words from the Bridge session (e.g., bridge word *haner* and terminus word *hajer*). To determine whether vocabulary that individuals learned in the Bridge session transferred to the Terminus session, we analyzed the data by first assigning each terminus word for each participant to one of three categories based on how well their substitution neighbors were learned during the Bridge session. Items in the Known Neighbor category were neighbors of bridge words that an individual got 4–5 out of 5 letters correct in the prior session. The Partly-Known Neighbor category included neighbors of bridge words with a score between 1 and 3 letters correct, and the Unknown Neighbor category included neighbors of bridge words that got a score of 0 letters correct. Note that items were assigned to Bridge-Knowledge conditions individually for each participant based on their performance in the Bridge session, and thus categories

have an unbalanced number of items [see **Table 1**; $\chi^2(2) = 96.39$, $p < 0.001$].

The model included fixed effects of Similarity Condition (Familiar vs. Unfamiliar) and Bridge-Knowledge (Known vs. Unknown Neighbor, and Known vs. Partly-Known Neighbor) plus interactions, as well as random intercepts for participant, word form, and word meaning, by-participant and by-form random slopes for Bridge-Knowledge, and by-meaning random slopes for Bridge-Knowledge and Similarity. Imageability, familiarity, age of acquisition, and word frequency of terminus words' English translations were entered as covariates.

We found a significant interaction between Similarity and Bridge-Knowledge [Known vs. Partly-Known contrast, *Estimate* = 15.02, *SE* = 6.29, 95% CI (2.69, 27.34), $t(66.5) = 2.39$, $p = 0.019$], but not the [Known vs. Unknown contrast; *Estimate* = 6.25, *SE* = 5.25, 95% CI (−4.04, 16.54), $t(96.6) = 1.19$, $p = 0.237$] and a main effect of Bridge-Knowledge [Known vs. Unknown contrast, *Estimate* = −11.69, *SE* = 3.12, 95% CI (−17.82, −5.57), $t(96.8) = −3.74$, $p < 0.001$; Known vs. Partly-Known contrast, *Estimate* = −11.35, *SE* = 3.45, 95% CI (−18.11, −4.6), $t(78.9)$

TABLE 1 | Percentage of terminus words with known, partly-known, and unknown neighbors.

	Known neighbors	Partly-known neighbors	Unknown neighbors
Familiar condition	29.2%	17.9%	53.1%
Unfamiliar condition	11.0%	17.4%	71.6%

Percentages reflect the average proportion of Terminus words with Known, Partly-Known, and Unknown Bridge word neighbors in the Familiar and Unfamiliar conditions. Each Terminus word's (e.g., hajer) category was determined on a participant-by-participant basis depending on how well its orthographically-related Bridge word (e.g., haner) was learned in the previous test session. A Terminus word was categorized as having a Known Neighbor if the participant correctly recalled 4–5 out of 5 letters of its corresponding Bridge word, as Partly-Known if they correctly recalled 1–3 letters, and as Unknown if they correctly recalled 0 letters. To determine the overall distribution of Terminus words in each category, the percentage of words with Known, Partly-Known, and Unknown neighbors was calculated for each participant and then averaged across participants within the Familiar and Unfamiliar conditions.

$= -3.30$, $p = 0.001$] (**Figure 2**). Follow-up comparisons on the model's predicted marginal means revealed that accuracy for Known Neighbor words in the Unfamiliar condition $M = 30.75$ [$SE = 5.37$, 95% CI (20.0, 41.5)] was higher than for both Partly-Known Neighbors $M = 12.68$ [$SE = 4.35$, 95% CI (3.8, 21.5)], $t(45.4) = 3.49$, $p = 0.003$ or Unknown Neighbors $M = 16.25$ [$SE = 3.2$, 95% CI (9.8, 22.7)], $t(50.0) = 3.18$, $p = 0.007$. In contrast, accuracy for Known Neighbor words in the Familiar condition $M = 27.60$, [$SE = 4.65$, 95% CI (18.2, 37.0)] did not differ from either Partly-Known Neighbors $M = 24.55$, [$SE = 4.81$, 95% CI (14.8, 34.3)], $t(24.4) = 0.67$, $p = 0.781$ or Unknown Neighbors $M = 19.36$, [$SE = 3.61$, 95% CI (12.1, 26.7)], $t(31.0) = 2.16$, $p = 0.094$[2]. Partly-Known Neighbors did not differ from Unknown Neighbors in either condition. These results show that learning a word in the Bridge session increased one's chances of learning its neighbor in the Terminus session, providing evidence that similarity to previously-learned novel words benefits later vocabulary acquisition. Critically, similarity to previously-learned words influenced the types of words that people learned in the Unfamiliar condition more than in the Familiar condition. The significant difference between Known Neighbors and both Partly-Known and Unknown Neighbors, but not between Partly-Known and Unknown Neighbors further suggests that complete acquisition of a bridge word was necessary for participants in the Unfamiliar condition to benefit from similarity to previously-learned words. Partially learning a bridge word did not result in any differences between the Familiar and Unfamiliar conditions.
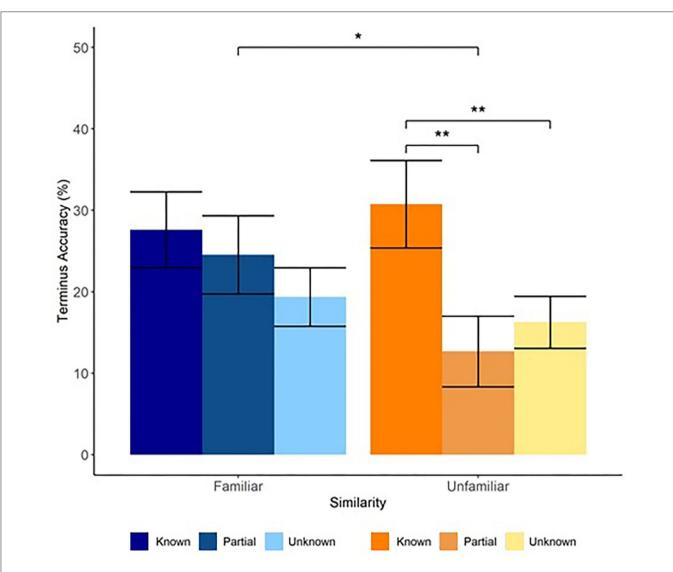
DISCUSSION

The goal of the present study was to determine how wordform similarity to the native language (as determined by bigram

---

[2]Similar results were obtained when Bridge word accuracy was instead entered as a continuous variable, showing a significant effect of Bridge-Knowledge in the Unfamiliar condition [*Estimate* $= 10.65$, $SE = 4.56$, $t(49.8) = 2.34$, $p = 0.024$], but a marginal effect in the Familiar condition [*Estimate* $= 7.71$, $SE = 3.95$, $t(83.8) = 1.95$, $p = 0.054$].

probability) influences acquisition of non-native vocabulary and sensitivity to non-native sublexical regularities. We found that although native language similarity provides short-term benefits, it can reduce reliance on non-native patterns during subsequent learning. Through continued use of an L2, the learner recognizes new patterns that determine how letters or sounds can combine to form words, and how words combine to form sentences. This process of extracting new patterns is also important for establishing continuous vocabulary learning, by ensuring that new words are accurately perceived and encoded in memory. Advanced L2 learners have been shown to benefit from L2 similarity during word learning (Stamer and Vitevitch, 2012; Ma et al., 2015), and in the current study, we found that similarity to other non-native words can also affect the earliest stages of vocabulary acquisition. Specifically, learning a word in the first session increased the likelihood that a similar word would be acquired in the subsequent session. Notably, while words that resembled the L1 were easier to learn at first, they had less of an influence on subsequent word learning. These results demonstrate the important roles of the native language, the burgeoning non-native vocabulary, and their interactions on new word learning.

Because of the way the new vocabulary in our study was designed, each word in the bridge session had a single substitution neighbor in the subsequent terminus session. These bridge-terminus word pairs allowed us to assess differences in word learning based on whether or not the learner already knew a similar word. Importantly, this is based not on intrinsic properties of the words, but instead on learners' idiosyncratic knowledge of patterns in the new word lists. Given the self-directed nature of the training session, the effect of similarity to previously-learned words that we observed may reflect how attention and study time were allocated to new words. Because overall accuracy did not improve between bridge and terminus sessions, the observed advantage for terminus words with already-acquired bridge neighbors comes at the expense of words with unlearned neighbors, consistent with prior self-directed word learning paradigms (Bardhan, 2010). The relative disadvantage for words with partly-learned neighbors may additionally result from the confusion that can occur between formally similar L2 words (e.g., the German words *Schafe,* meaning "sheep" and *schaffen,* meaning "create"; Laufer, 1988, 1989). Laufer (1988) conjectures that these types of "synform errors" may result from weak or unstable representations of L2 words in memory that could impair the learner's ability to distinguish between them or correctly map them to their corresponding meanings. Transfer from previously-learned wordforms may therefore have contrasting effects on subsequent learning depending on how well the initial words were learned, with facilitation from robustly encoded exemplars but interference from more unstable representations.

Notably, learners in the Familiar and Unfamiliar conditions differed in how much similarity to previously-learned words affected their continued learning. Even though bridge words in the first session were learned twice as well in the Familiar condition compared to the Unfamiliar condition, the effect of learning similar words in the terminus session was nearly twice as

FIGURE 2 | Similarity to previously-learned novel words influences later acquisition. Word learning in the Bridge session affected the likelihood of learning its orthographic neighbor in the Terminus session. The effect of prior novel word learning was moderated by similarity to the native language. Accuracy in the Unfamiliar condition was higher for Known Neighbor words (dark orange) than for Partly-Known Neighbor words (orange) and Unknown Neighbor words (light orange); accuracy for Partly-Known and Unknown Neighbor words did not differ from each other. Accuracy in the Familiar condition did not differ between Known Neighbor words (dark blue), Partly-Known Neighbor words (blue), and Unknown Neighbor words (light blue). Error bars represent standard error (by participants). $*p < 0.05$; $**p < 0.01$.

large for learners in the *Unfamiliar* condition. In the Unfamiliar condition, terminus words with learned bridge word neighbors were recalled with 2.65 times greater accuracy than words with unlearned bridge neighbors, compared to only a 1.65 times advantage in the Familiar condition. These terminus words in the second session were carefully designed to have equally low English similarity in both conditions, ensuring that this terminus word difference was due to effects of similarity to other non-native words, without confounding native and non-native word similarity. Together, these results indicate that although native language similarity provided an early benefit for word learning, it reduced the benefit of similarity to previously-learned non-native words in continued study.

Part of the task of learning a second language and achieving lexical competence involves building a foundation of L2 knowledge and a network of connections among L2 words and their meanings (Ellis and Beaton, 1993; De Groot, 2011), which can enhance the automaticity of L2 processing and minimize reliance on, and interference from, L1 knowledge (MacWhinney, 1997; Jiang, 2000). Connectionist models of bilingual language processing suggest that language selection and control can be accomplished over time via Hebbian learning and self-organizing representations that naturally cluster in language-specific ways due to greater feature overlap and co-activation of words within-languages, than across languages (e.g., Shook and Marian's, 2013 *BLINCS* model). Such a system could allow bilinguals to rely on bottom-up inputs, such as orthographic or phonological features, to selectively activate the appropriate language based on learned regularities within each language. For instance, language-specific sublexical cues, such as letter and bigram frequencies,

can reduce the activation of cross-linguistic primes (Casaponsa and Duñabeitia, 2016; Dubey et al., 2018), and bilinguals can rely on language membership cues to guide lexical access (Grainger and Beauvillain, 1987; Vaid and Frenck-Mestre, 2002; Casaponsa et al., 2014) and speech production (Oganian et al., 2015). Participants in the current study who learned Familiar bridge words did not have orthographic cues that could reliably indicate language membership prior to lexical processing, which could have increased the activation of English representations relative to participants in the Unfamiliar condition. This may have stalled the process of linking new words into a coherent L2, interfering with transfer between the bridge and terminus words. In contrast, learners in the Unfamiliar condition were acquiring vocabulary that was unambiguously distinct from English. This distinction appears to be helpful in promoting extraction of non-native patterns to be used during new word learning.

The fact that the Familiar and Unfamiliar conditions did not differ in overall terminus word accuracy, however, may indicate that the two groups made use of different strategies or could have differed in other meaningful ways, such as in motivation, which has been shown to benefit word learning (Dörnyei and Ushioda, 2009; Dörnyei, 2019). For instance, the relative ease of learning bridge words that were similar to the L1 could have reduced motivation and effort in the Familiar condition, particularly during the second session when the task was unexpectedly more difficult. This could have elicited shallower processing of the terminus words during the word learning phase and consequently, reduced transfer from known bridge words. A complimentary interpretation would be that learners in the Unfamiliar condition benefited from the "desirable difficulties" (Bjork and Bjork, 2011) associated with learning more challenging bridge words. In language learning, retention in long term memory is generally improved when learning requires a greater depth of processing and involvement (Craik and Lockhart, 1972; Laufer and Hulstijn, 2001; Rice and Tokowicz, 2020), which can be instigated by material presented in a more difficult context (Schneider et al., 2002; Bjork and Kroll, 2015). Examples of desirable difficulties include repeated testing in place of passive study, or interleaving blocks of different word lists rather than blocked study (Schneider et al., 2002; Bjork and Kroll, 2015; Marecka et al., 2021). Our results suggest that difficulties caused by properties of the words themselves may also be targets for increasing long-term learning. Future research incorporating measures of motivation and/or manipulations of task engagement (e.g., through game-like formats, De Vos et al., 2019; Yang et al., 2020; see Derakhshan and Khatir, 2015 for review) could help elucidate the potential role of affective variables in determining the impact of cross-linguistic influence on transfer between non-native vocabulary.

In conclusion, we found that new vocabulary learning is affected by both similarity to one's native language and similarity to other newly learned words. Whereas native language similarity has a beneficial effect early on, it may decrease sensitivity to non-native word patterns that support later learning. This suggests that cross-linguistic influence is modulated by interactions

between existing native and non-native word knowledge, and that initial similarity to the native language can have dynamically changing consequences over the course of novel word learning. This is because the words that one successfully learns early-on can influence the words that one acquires later, by driving attention toward new words that look more like already acquired ones. This suggests that cross-linguistic influences on initial vocabulary learning could potentially have cascading effects on the makeup of one's later vocabulary. Overall, these results demonstrate the complex relationship between native and non-native vocabulary, where similarity can have variable consequences for learning.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Northwestern University Institutional Review Board. The participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Conceptualization and design: VM, JB, and AB. Data collection: JB and AB. Statistical analysis: JB and SH. Writing—original draft preparation: VM and JB. Writing—review and editing: SH, VM, and AB. Project administration and funding acquisition: VM. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2021.651506/full#supplementary-material

REFERENCES

Amengual, M. (2012). Interlingual influence in bilingual speech: cognate status effect in a continuum of bilingualism. *Bilingualism* 15, 517–530. doi: 10.1017/S1366728911000460

Amengual, M. (2016). Cross-linguistic influence in the bilingual mental lexicon: evidence of cognate effects in the phonetic production and processing of a vowel contrast. *Front. Psychol.* 7:617. doi: 10.3389/fpsyg.2016.00617

Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge: Harvard University Press.

Atkinson, R. C., and Raugh, M. R. (1975). An application of the mnemonic keyword method to the acquisition of a Russian vocabulary. *J. Exp. Psychol.* 1, 126–133. doi: 10.1037/0278-7393.1.2.126

Bardhan, N. (2010). *Adults' self-directed learning of an artificial lexicon: the dynamics of neighborhood reorganization* (Ph.D. Thesis). University of Rochester, Rochester, NY, USA.

Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001

Bartolotti, J., and Marian, V. (2014). Wordlikeness and novel word learning. *Proc. Annu. Meet. Cogn. Sci. Soc.* 36, 146–151. Available online at: https://escholarship.org/uc/item/2j8652jh

Bartolotti, J., and Marian, V. (2017). Orthographic knowledge and lexical form influence vocabulary learning. *Appl. Psycholinguist.* 38, 427–456. doi: 10.1017/S0142716416000242

Bartolotti, J., Marian, V., Schroeder, S. R., and Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Front. Psychol.* 2:324. doi: 10.3389/fpsyg.2011.00324

Bates, D. M., Machler, M., Bolker, B. M., and Walker, S. C. (2014). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67:121811. doi: 10.18637/jss.v067.i01

Bijeljac-Babic, R., Biardeau, A., and Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Mem. Cogn.* 25, 447–457. doi: 10.3758/BF03201121

Bjork, E. L., and Bjork, R. A. (2011). "Making things hard on yourself, but in a good way: creating desirable difficulties to enhance learning," in *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society, 2nd Edn*, eds M. A. Gernsbacher, and J. Pomerantz (New York, NY: Worth), 59–68.

Bjork, R. A., and Kroll, J. F. (2015). Desireable difficulties in vocabulary learning. *Am. J. Psychol.* 128, 241–252. doi: 10.5406/amerjpsyc.128.2.0241

Braun, M., Hutzler, F., Ziegler, J. C., Dambacher, M., and Jacobs, A. M. (2009). Pseudohomophone effects provide evidence of early lexico-phonological processing in visual word recognition. *Hum. Brain Mapp.* 30, 1977–1989. doi: 10.1002/hbm.20643

Brooks-Lewis, K. A. (2009). Adult learners' perceptions of the incorporation of their L1 in foreign language teaching and learning. *Appl. Linguist.* 30, 216–235. doi: 10.1093/applin/amn051

Brysbaert, M., and New, B. (2009). Moving beyond Kucera and Francis: a critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behav. Res. Methods* 41, 977–990. doi: 10.3758/BRM.41.4.977

Brysbaert, M., Van Dyck, G., and Van de Poel, M. (1999). Visual word recognition in bilinguals: evidence from masked phonological priming. *J. Exp. Psychol.* 25, 137–148. doi: 10.1037/0096-1523.25.1.137

Brysbaert, M., and Van Wijnendaele, I. (2003). The importance of phonological coding in visual word recognition: further evidence from second-language processing. *Psychol. Belg.* 43, 249–258. doi: 10.5334/pb.1011

Casaponsa, A., Carreiras, M., and Duñabeitia, J. A. (2014). Discriminating languages in bilingual contexts: the impact of orthographic markedness. *Front. Psychol.* 5, 424. doi: 10.3389/fpsyg.2014.00424

Casaponsa, A., and Duñabeitia, J. A. (2016). Lexical organization of language-ambiguous and language-specific words in bilinguals. *Q. J. Exp. Psychol.* 69, 589–604. doi: 10.1080/17470218.2015.1064977

Casaponsa, A., Thierry, G., and Duñabeitia, J. A. (2020). The role of orthotactics in language switching: an ERP investigation using masked language priming. *Brain Sci.* 10:22. doi: 10.3390/brainsci10010022

Cheung, H. (1996). Nonword span as a unique predictor of second-language vocabulary language. *Dev. Psychol.* 32, 867–873. doi: 10.1037/0012-1649.32.5.867

Comesaña, M., Ferré, P., Romero, J., Guasch, M., Soares, A. P., and García-Chico, T. (2015). Facilitative effect of cognate words vanishes when reducing the orthographic overlap: the role of stimuli list composition. *J. Exp. Psychol.* 41, 614–635. doi: 10.1037/xlm0000065

Comesaña, M., Perea, M., Piñeiro, A., and Fraga, I. (2009). Vocabulary teaching strategies and conceptual representations of words in L2 in children: Evidence with novice learners. *J. Exp. Child Psychol.* 104, 22–33. doi: 10.1016/j.jecp.2008.10.004

Comesaña, M., Soares, A. P., Sánchez-Casas, R., and Lima, C. (2012). Lexical and semantic representations in the acquisition of L2 cognate and non-cognate words: evidence from two learning methods in children. *Br. J. Psychol.* 103, 378–392. doi: 10.1111/j.2044-8295.2011.02080.x

Craik, F. I. M., and Lockhart, R. S. (1972). Levels of processing: a framework for memory research. *J. Verbal Learn. Verbal Behav.* 11, 671–684. doi: 10.1016/S0022-5371(72)80001-X

Cunningham, T. H., and Graham, C. (2000). Increasing native English vocabulary recognition through Spanish immersion: cognate transfer from foreign to first language. *J. Educ. Psychol.* 92, 37–49. doi: 10.1037/0022-0663.92.1.37

De Groot, A. M. (2006). Effects of stimulus characteristics and background music on foreign language vocabulary learning and forgetting. *Lang. Learn.* 56, 463–506. doi: 10.1111/j.1467-9922.2006.00374.x

De Groot, A. M., and Keijzer, R. (2000). What is hard to learn is easy to forget: the roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Lang. Learn.* 50, 1–56. doi: 10.1111/0023-8333.00110

De Groot, A. M. B. (2011). *Language and Cognition in Bilinguals and Multilinguals: An Introduction*. New York, NY: Psychology Press. doi: 10.4324/9780203841228

De Vos, J. F., Schriefers, H., and Lemhöfer, K. (2019). Noticing vocabulary holes aids incidental language word learning: an experimental study. *Bilingualism* 22, 500–515. doi: 10.1017/S1366728918000020

Derakhshan, A., and Khatir, E. D. (2015). The effects of using games on English vocabulary learning. *J. Appl. Linguist. Lang. Res.* 2, 39–47. doi: 10.36892/ijlls.v1i1.22

Dixon, L. Q., Zhao, J., Shin, J. Y., Wu, S., Su, J. H., Burgess-Brigham, R., et al. (2012). What we know about second language acquisition: a synthesis from four perspectives. *Rev. Educ. Res.* 82, 5–60. doi: 10.3102/0034654311433587

Dörnyei, Z. (2019). "Task motivation. What makes an L2 task engaging?," in *Researching L2 Task Performance and Pedagogy: In Honour of Peter Skehan*, eds Z. E. Wen, and M. J. Ahmadian (Amsterdam: John Benjamins), 53–66.

Dörnyei, Z., and Ushioda, E. (2009). *Motivation, Language Identity and the L2 Self*. Bristol: Multilingual Matters. doi: 10.21832/9781847691293

Dressler, C., Carlo, M. S., Snow, C. E., August, D., and White, C. E. (2011). Spanish-speaking students' use of cognate knowledge to infer the meaning of English words. *Bilingualism* 14, 243–255. doi: 10.1017/S1366728910000519

Dubey, N., Witzel, N., and Witzel, J. (2018). Script differences and masked translation priming: evidence from Hindi-English bilinguals. *Q. J. Exp. Psychol.* 71, 2421–2438. doi: 10.1177/1747021817743241

Duddington, J. (2012). *eSpeak*. Available online at: http://espeak.sourceforge.net (accessed August 26, 2016).

Ecke, P. (2015). Parasitic vocabulary acquisition, cross-linguistic influence, and lexical retrieval in multilinguals. *Bilingualism* 18, 145–162. doi: 10.1017/S1366728913000722

Eckman, F. R. (2004). From phonemic differences to constraint rankings: research on second language phonology. *Stud. Second Lang. Acquisit.* 26, 513–549. doi: 10.1017/S027226310404001X

Ellis, N. C. (2008). *Usage-Based and Form-Focused Language Acquisition: The Associative Learning of Constructions, Learned Attention, and the Limited L2 Endstate*. London: Routledge/Taylor and Francis Group.

Ellis, N. C., and Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Lang. Learn.* 43, 559–617. doi: 10.1111/j.1467-1770.1993.tb00627.x

Ellis, N. C., and Sinclair, S. (1996). Working memory in the acquisition of vocabulary and syntax: putting language in good order. *Q. J. Exp. Psychol.* 49, 234–250. doi: 10.1080/713755604

Ferrand, L., and Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bull. Psychonom. Soc.* 31, 119–122. doi: 10.3758/BF03334157

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: evidence for the effect of equivalence classification. *J. Phon.* 15, 47–65. doi: 10.1016/S0095-4470(19)30537-6

Gathercole, S. E., and Masoura, E. V. (2005). Contrasting contributions of phonological short-term memory and long-term knowledge to vocabulary learning in a foreign language. *Memory* 13, 422–429. doi: 10.1080/09658210344000323

Goldrick, M., Runnqvist, E., and Costa, A. (2014). Language switching makes pronunciation less nativelike. *Psychol. Sci.* 25, 1031–1036. doi: 10.1177/0956797613520014

Grainger, J., and Beauvillain, C. (1987). Language blocking and lexical access in bilinguals. *Q. J. Exp. Psychol.* 39, 295–319. doi: 10.1080/14640748708401788

Grainger, J., and Ferrand, L. (1994). Phonology and orthography in visual word recognition: effects of masked homophone primes. *J. Mem. Lang.* 33, 218–233. doi: 10.1006/jmla.1994.1011

Grainger, J., Kiyonaga, K., and Holcomb, P. J. (2006). The time course of orthographic and phonological code activation. *Psychol. Sci.* 17, 1021–1026. doi: 10.1111/j.1467-9280.2006.01821.x

Green, P., and Macleod, C. J. (2016). SIMR: an R package for power analysis of generalized linear mixed models by simulation. *Methods Ecol. Evol.* 7, 493–498. doi: 10.1111/2041-210X.12504

Hakuta, K. (1982). Interaction between particles and word order in the comprehension and production of simple sentences in Japanese children. *Dev. Psychol.* 18, 62–76. doi: 10.1037/0012-1649.18.1.62

Hamada, M., and Koda, K. (2008). Influence of first language orthographic experience on second language decoding and word learning. *Lang. Learn.* 58, 1–31. doi: 10.1111/j.1467-9922.2007.00433.x

Hancin-Bhatt, B., and Nagy, W. (1994). Lexical transfer and second language morphological development. *Appl. PsychoLinguist.* 15, 289–310. doi: 10.1017/S0142716400065905

Helms-Park, R., and Perhan, Z. (2016). The role of explicit instruction in cross-script cognate recognition: The case of Ukrainian-speaking EAP learners. *J. English Acad.* 21, 17–33. doi: 10.1016/j.jeap.2015.08.005

Hermann, F. (2003). Differential effects of reading and memorization of paired associates on vocabulary acquisition in adult learners of English as a second language. *TESL-EJ* 7, 1–16.

Holcomb, P. J., Grainger, J., and O'rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *J. Cogn. NeuroSci.* 14, 938–950. doi: 10.1162/089892902760191153

Horst, M., Cobb, T., and Meara, P. (1998). Beyond a clockwork orange: acquiring second language vocabulary through reading. *Read. Foreign Lang.* 11, 207–223.

Hulstijn, J. H. (1997). "Mnemonic methods in foreign language vocabulary learning: theoretical considerations and pedagogical implications," in *Second Language Vocabulary Acquisition: A Rationale for Pedagogy,* eds J. Coady, and E. Huckin (Cambridge: Cambridge University Press), 203–224. doi: 10.1017/CBO9781139524643.015

Jarvis, S., and Pavlenko, A. (2008). *Crosslinguistic Influence in Language and Cognition*. New York, NY: Routledge. doi: 10.4324/9780203935927

Jiang, N. (2000). Lexical representation and development in a second language. *Appl. Linguist.* 21, 47–77. doi: 10.1093/applin/21.1.47

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., and Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: an fMRI study. *Brain Lang.* 127, 46–54. doi: 10.1016/j.bandl.2012.11.007

Kaushanskaya, M., and Marian, V. (2009). Bilingualism reduces native-language interference during novel-word learning. *J. Exp. Psychol.* 35, 829–835. doi: 10.1037/a0015275

Kellerman, E. (1978). *Giving Learners a Break: Native Language Intuitions as a Source of Predictions About Transferability. Working Papers on Bilingualism.* Toronto, ON: Ontario Institute for Studies in Education, 59–92.

Koda, K. (1989). The effects of transferred vocabulary knowledge on the development of L2 reading proficiency. *Foreign Lang. Ann.* 22, 529–540. doi: 10.1111/j.1944-9720.1989.tb02780.x

Koda, K. (1990). The use of L1 reading strategies in L2 reading: effects of L1 orthographic structures on L2 phonological recoding strategies. *Stud. Second Lang. Acquisit.* 12, 393–410. doi: 10.1017/S0272263100009499

Koda, K. (1993). Transferred L1 strategies and L2 syntactic structure in L2 sentence comprehension. *Modern Lang. J.* 77, 490–500. doi: 10.1111/j.1540-4781.1993.tb01997.x

Koda, K. (1998). The role of phonemic awareness in second language reading. *Second Lang. Res.* 14, 194–215. doi: 10.1191/026765898676398460

Koda, K., and Zehler, A. M. (2008). *Learning to Read Across Languages: Cross-Linguistic Relationships in First-and Second-Language Literacy Development*. New York, NY: Routledge. doi: 10.4324/9780203935668

Laufer, B. (1988). The concept of 'synforms' (similar lexical forms) in vocabulary acquisition. *Lang. Educ.* 2, 113–132. doi: 10.1080/09500788809541228

Laufer, B. (1989). A factor of difficulty in vocabulary learning: deceptive transparency. *AILA Rev.* 6, 10–20.

Laufer, B., and Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: a case for contrastive analysis and translation. *Appl. Linguist.* 29, 694–716. doi: 10.1093/applin/amn018

Laufer, B., and Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: the construct of task-induced involvement. *Appl. Linguist.* 22, 1–26. doi: 10.1093/applin/22.1.1

Laufer, B., and Shmueli, K. (1997). Memorizing new words: does teaching have anything to do with it? *RELC J.* 28, 89–108. doi: 10.1177/003368829702800106

Lin, A. M. Y. (2015). Conceptualising the potential role of L1 in CLIL. *Lang. Cult. Curriculum* 28, 74–89. doi: 10.1080/07908318.2014.1000926

Liskin-Gasparro, J. E. (1982). *ETS oral Proficiency Testing Manual*. Princeton, NJ: Educational Testing Service.

Lotto, L., and de Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Lang. Learn.* 48, 31–69. doi: 10.1111/1467-9922.00032

Ma, T., Chen, B., Lu, C., and Dunlap, S. (2015). Proficiency and sentence constraint effects on second language word learning. *Acta Psychol.* 159, 116–122. doi: 10.1016/j.actpsy.2015.05.014

MacWhinney, B. (1997). "Second language acquisition and the competition model," in *Tutorials in Bilingualism: Psycholinguistic Perspectives,* eds A. M. B. de Groot, and J. F. Kroll (Mahwah, NJ: Lawrence Erlbaum Associates Publishers), 113–142.

MacWhinney, B., and Bates, E. (1989). *The Crosslinguistic Study of Sentence Processing*. Cambridge: Cambridge University Press.

Majerus, S., Poncelet, M., Van der Linden, M., and Weekes, B. S. (2008). Lexical learning in bilingual adults: the relative importance of short-term memory for serial order and phonological knowledge. *Cognition* 107, 395–419. doi: 10.1016/j.cognition.2007.10.003

Marecka, M., Szewczyk, J., Otwinowska, A., Durlik, J., Foryś-Nogala, M., Kutyłowska, K., et al. (2021). False friends or real friends? False cognates show advantage in word form learning. *Cognition* 206:104477. doi: 10.1016/j.cognition.2020.104477

Marian, V., Bartolotti, J., Chabal, S., and Shook, A. (2012). CLEARPOND: cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PloS ONE* 7:e43230. doi: 10.1371/journal.pone.0043230

Marian, V., Blumenfeld, H. K., and Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): assessing language profiles in bilinguals and multilinguals. *J. Speech Lang. Hear. Res.* 50, 940–967. doi: 10.1044/1092-4388(2007/067)

McLaughlin, J., Osterhout, L., and Kim, A. (2004). Neural correlates of second-language word learning: minimal instruction produces rapid change. *Nat. NeuroSci.* 7, 703–704. doi: 10.1038/nn1264

Meade, G., Midgley, K. J., Dijkstra, T., and Holcomb, P. J. (2018). Cross-language neighborhood effects in learners indicative of an integrated lexicon. *J. Cogn. NeuroSci.* 30, 70–85. doi: 10.1162/jocn_a_01184

Meara, P. (1980). Vocabulary acquisition: a neglected aspect of language learning. *Lang. Teach. Linguist.* 13, 221–246. doi: 10.1017/S0261444800008879

Melby-Lervåg, M., and Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: a meta-analysis of the correlational evidence. *J. Res. Read.* 34, 114–135. doi: 10.1111/j.1467-9817.2010.01477.x

Mirman, D., Magnuson, J. S., Graf Estes, K., and Dixon, J. A. (2008). The link between statistical segmentation and word learning in adults. *Cognition* 108, 271–280. doi: 10.1016/j.cognition.2008.02.003

Mori, Y. (1998). Effects of first language and phonological accessibility on kanji recognition. *Modern Lang. J.* 82, 69–82. doi: 10.1111/j.1540-4781.1998.tb02595.x

Nation, I. P. (1982). Beginning to learn foreign vocabulary: a review of the research. *RELC J.* 13, 14–36. doi: 10.1177/003368882201300102

Odlin, T. (1989). *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge: Cambridge University Press.

Oganian, Y., Conrad, M., Aryani, A., Heekeren, H. R., and Spalek, K. (2015). Interplay of bigram frequency and orthographic neighborhood statistics in language membership decision. *Bilingualism* 19, 578–596. doi: 10.1017/S1366728915000292

Ortega, L. (2013). SLA for the 21st century: disciplinary progress, transdisciplinary relevance, and the bi/multilingual turn. *Lang. Learn.* 63, 1–24. doi: 10.1111/j.1467-9922.2012.00735.x

Osterhout, L., McLaughlin, J., Pitkänen, I., Frenck-Mestre, C., and Molinaro, N. (2006). Novice learners, longitudinal designs, and event-related potentials: a means for exploring the neurocognition of second language processing. *Lang. Learn.* 56, 199–230. doi: 10.1111/j.1467-9922.2006.00361.x

Otwinowska, A. (2015). *Cognate Vocabulary in Language Acquisition and Use*. Bristol: Multilingual Matters. doi: 10.21832/9781783094394

Otwinowska, A., Foryś-Nogala, M., Kobosko, W., and Szewczyk, J. (2020). Learning orthographic cognates and non-cognates in the classroom: does awareness of cross-linguistic similarity matter? *Lang. Learn.* 70, 685–731. doi: 10.1111/lang.12390

Otwinowska, A., and Szewczyk, J. M. (2019). The more similar the better? Factors in learning cognates, false cognates and non-cognate words. *Int. J. Bilingual Educ. Bilingualism* 22, 974–991. doi: 10.1080/13670050.2017.1325834

Paivio, A., and Desrochers, A. (1981). Mnemonic techniques in second-language learning. *J. Educ. Psychol.* 73, 780–795. doi: 10.1037/0022-0663.73.6.780

Papagno, C., Valentine, T., and Baddeley, A. (1991). Phonological short-term memory and foreign-language vocabulary learning. *J. Mem. Lang.* 30, 331–347. doi: 10.1016/0749-596X(91)90040-Q

Perfetti, C. A., and Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: evidence from backward masking and priming. *J. Mem. Lang.* 30, 473–485. doi: 10.1016/0749-596X(91)90017-E

Pickering, M. (1982). Context-free and context-dependent vocabulary learning: an experiment. *System* 10, 79–83. doi: 10.1016/0346-251X(81)90070-1

Prince, P. (1996). Second language vocabulary learning: the role of context versus translations as a function of proficiency. *Modern Lang. J.* 80, 478–493. doi: 10.1111/j.1540-4781.1996.tb05468.x

PsychCorp (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Harcourt Assessment.

Puimège, E., and Peters, E. (2019). Learner's English vocabulary knowledge prior to formal instruction: the role of word-related and learner-related variables. *Lang. Learn.* 69, 943–977. doi: 10.1111/lang.12364

Pulido, D. (2003). Modeling the role of second language proficiency and topic familiarity in second language incidental vocabulary acquisition through reading. *Lang. Learn.* 53, 233–284. doi: 10.1111/1467-9922.00217

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: https://www.r-project.org (accessed March 22, 2021).

Rice, C. A., and Tokowicz, N. (2020). State of the scholarship: a review of laboratory studies of adult second language vocabulary training. *Stud. Second Lang. Acquisit.* 42, 439–470. doi: 10.1017/S0272263119000500

Ringbom, H. (2007). Actual, perceived and assumed cross-linguistic similarities in foreign language learning. *AFinLAn vuosikirja*. doi: 10.21832/9781853599361

Ringbom, H., and Jarvis, S. (2011). "The importance of cross-linguistic similarity in foreign language learning," in *The Handbook of Language Teaching*, eds Long M. H., and Doughty C. J. (West Sussex, Wiley-Blackwell), 106–118. doi: 10.1002/9781444315783.ch7

Roberts, J. A., Pollock, K. E., Krakow, R., Price, J., Fulmer, K. C., and Wang, P. P. (2005). Language development in preschool-age children adopted from China. *J. Speech Lang. Hear. Res.* 48:93. doi: 10.1044/1092-4388(2005/008)

Rogers, J., Webb, S., and Nakata, T. (2015). Do the cognacy characteristics of loanwords make them more easily learned than noncognates? *Lang. Teach. Rese.* 19, 9–27. doi: 10.1177/1362168814541752

Roodenrys, S., and Hinton, M. (2002). Sublexical or lexical effects on serial recall of nonwords? *J. Exp. Psychol.* 28, 29–33. doi: 10.1037/0278-7393.28.1.29

Saffran, J. R., Johnson, E. K., Aslin, R. N., and Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition* 70, 27–52. doi: 10.1016/S0010-0277(98)00075-4

Schneider, V. I., Healy, A. F., and Bourne, L. E. (2002). What is learned under difficult conditions is hard to forget: contextual interference effects in foreign vocabulary acquisition, retention, and transfer. *J. Mem. Lang.* 46, 419–440. doi: 10.1006/jmla.2001.2813

Service, E., and Craik, F. I. M. (1993). Differences between young and older adults in learning a foreign vocabulary. *J. Mem. Lang.* 32, 608–623. doi: 10.1006/jmla.1993.1031

Shirai, Y. (1992). Conditions on transfer: a connectionist approach. *Appl. Linguist.* 3, 91–120.

Shook, A., and Marian, V. (2013). The bilingual language interaction network for comprehension of speech. *Bilingualism* 16, 304–324. doi: 10.1017/S1366728912000466

Stadthagen-Gonzalez, H., and Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Beh. Res. Methods* 38, 598–605. doi: 10.3758/BF03193891

Stamer, M. K., and Vitevitch, M. S. (2012). Phonological similarity influences word learning in adults learning Spanish as a foreign language. *Bilingualism* 15, 490–502. doi: 10.1017/S1366728911000216

Storkel, H. L., Armbrüster, J., and Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *J. Speech Lang. Hear. Res.* 49, 1175–1192. doi: 10.1044/1092-4388(2006/085)

Storkel, H. L., and Maekawa, J. (2005). A comparison of homonym and novel word learning: The role of phonotactic probability and word frequency. *J. Child. Lang.* 32, 827–853. doi: 10.1017/S0305000905007099

Tekmen, E. A. F., and Daloglu, A. (2006). An investigation of incidental vocabulary acquisition in relation to learner proficiency level and word frequency. *Foreign Lang. Ann.* 39, 220–243. doi: 10.1111/j.1944-9720.2006.tb02263.x

Tonzar, C., Lotto, A., and Job, R. (2009). L2 vocabulary acquisition in children: effects of learning method and cognate status. *Lang. Learn.* 59, 623–646. doi: 10.1111/j.1467-9922.2009.00519.x

Tréville, M. C. (1996). Lexical learning and reading in L2 at the beginner level: the advantage of cognates. *Can. Modern Lang. Rev.* 53, 173–190. doi: 10.3138/cmlr.53.1.173

Vaid, J., and Frenck-Mestre, C. (2002). Do orthographic cues aid language recognition? A laterality study with French–English bilinguals. *Brain Lang.* 82, 47–53. doi: 10.1016/S0093-934X(02)00008-1

Van Hell, J. G., and Tanner, D. (2012). Second language proficiency and cross-language lexical activation. *Lang. Learn.* 62, 148–171. doi: 10.1111/j.1467-9922.2012.00710.x

Van Heuven, W. J. B., Mandera, P., Keuleers, E., and Brysbaert, M. (2014). SUBTLEX-UK: a new and improved word frequency database for British English. *Q. J. Exp. Psychol.* 67, 1176–1190. doi: 10.1080/17470218.2013.850521

Van Wijnendaele, I., and Brysbaert, M. (2002). Visual word recognition in bilinguals: phonological priming from the second to the first language. *J. Exp. Psychol.* 28, 616–627. doi: 10.1037/0096-1523.28.3.616

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Lang. Learn.* 61, 219–258. doi: 10.1111/j.1467-9922.2010.00593.x

Wang, T., and Saffran, J. R. (2014). Statistical learning of a tonal language: the influence of bilingualism and previous linguistic experience. *Front. Psychol.* 5:953. doi: 10.3389/fpsyg.2014.00953

Webb, S. (2007). Learning word pairs and glossed sentences: the effects of a single context on vocabulary knowledge. *Lang. Teach. Res.* 11, 63–81. doi: 10.1177/1362168806072463

Wechsler, D. (1997). *Wechsler Memory Scale III.* San Antonio, TX: Psychological Corporation.

Wrembel, M. (2011). "Cross-linguistic influence in third language acquisition of voice onset time," in *Proceedings of the 17th international congress of phonetic sciences*, eds W.-S. Lee, and E. Zee (Hong Kong: City University of Hong Kong), 2157–2160.

Yang, Q. F., Chang, S. C., Hwang, G. J., and Zou, D. (2020). Balancing cognitive complexity and gaming level: effects of a cognitive complexity-based competition game on EFL students' English vocabulary learning performance, anxiety and behaviors. *Comput. Educ.* 148:103808. doi: 10.1016/j.compedu.2020. 103808

Zahar, R., Cobb, T., and Spada, N. (2001). Acquiring vocabulary through reading: effects of frequency and contextual richness. *Can. Modern Lang. Rev.* 57, 541–572. doi: 10.3138/cmlr.57.4.541