

## Registered Report

**Cite this article:** Garrido-Pozú, J.J. (2024). Cross-linguistic effects of form overlap in aural recognition of Spanish–English cognates. *Bilingualism: Language and Cognition*, 27, 914–926. <https://doi.org/10.1017/S1366728924000270>

Received: 25 April 2023  
Revised: 7 March 2024  
Accepted: 8 March 2024  
First published online: 25 April 2024

### Keywords:

cognates; spoken word recognition; auditory processing; phonological overlap; orthographic overlap

### Corresponding author:

Juan J. Garrido Pozú;  
Email: [juan.garrido@furman.edu](mailto:juan.garrido@furman.edu)

## Abstract

This study investigated the effect of cross-linguistic overlap in L1 and L2 auditory recognition of Spanish–English cognates. The study examined the correlation between objective and subjective measures of overlap and analyzed how these measures predict patterns in auditory recognition. 62 Spanish-speaking learners of English and 63 English-speaking learners of Spanish completed two auditory lexical decision tasks in Spanish and English and a rating task, where they rated the perceived phonological similarity of cognates. The results revealed moderate correlations between subjective and objective measures of overlap. While orthographic overlap had no effect, increased phonological overlap facilitated recognition in L1 and L2 Spanish and English and had larger effects in L2 recognition. Perceived similarity was the best predictor among the measures of overlap. The findings support models suggesting that cross-linguistic co-activation is facilitated by increased form similarity and studies reporting modality dependent effects of cross-linguistic form overlap in lexical recognition.

## Introduction

Cognates are lexical items across languages that share meaning and high degrees of orthographic and phonological overlap. Cognates facilitate lexical processing compared to non-cognate words due to the increased cross-linguistic co-activation of cognate pairs (e.g., Dijkstra et al., 1999; Midgley et al., 2011; Peeters et al., 2013). Defining cognates in phonological terms is always a difficult task. Cognates are rarely phonologically identical due to phonological and phonetic differences across languages, and in many cases, cognates may differ in most sounds but still cause facilitation effects (Sherkina, 2003). Some cognates have complete orthographic overlap but are not phonologically identical, which shows that phonological and orthographic overlap do not have a direct positive linear relationship. Despite the abundant research on cognates, there is no consensus about the definition and classification of cognates. Many studies do not provide a detailed description of what they regard as ‘similar’. Available studies measure cross-linguistic orthographic/phonological overlap using different criteria, which makes an interpretation across different research findings less feasible.

The facilitation effect of cognates may depend on a variety of factors. Studies in visual word recognition report that cognate facilitation effects in bilinguals and L2 (i.e., second language) learners are more prominent with a higher degree of orthographic and phonological overlap (e.g., Carrasco-Ortiz et al., 2019; Comesaña et al., 2015; Dijkstra et al., 2010). However, differences in the degree of phonological overlap have also led to mixed findings, which suggests that orthographic and phonological overlap may affect bilingual lexical recognition in different ways (e.g., Carrasco-Ortiz et al., 2019; Comesaña et al., 2015; Dijkstra et al., 1999; Frances et al., 2021; Schwartz et al., 2007). Whereas some studies report facilitation effects of higher phonological overlap in visual word recognition (e.g., Dijkstra et al., 2010), other studies have reported inhibitory effects for higher phonological overlap, instead of facilitation effects, in visual recognition of cognates in language pairs that have similar phonetic repertoires (e.g., Comesaña et al., 2015; Dijkstra et al., 1999). Importantly, evidence of cognate similarity effects in spoken word recognition is still scarce. There is consistent evidence indicating that higher orthographic overlap benefits cognate recognition, but the effects of phonological overlap are still unclear.

The present study investigates cross-linguistic effects of cognates in spoken word recognition. The main goal of the present study is to examine how phonological overlap modulates auditory recognition of cognates in L2 learners of Spanish and English. In addition, the present study also provides a comparison of different subjective and objective measures of phonological similarity to test how they correlate with each other and how they relate to the patterns observed in auditory cognate recognition.

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

### Evidence of cross-linguistic facilitation effects of cognates

Cognate facilitation effects have been reported in a variety of studies in visual word recognition (e.g., Dijkstra et al., 2010), spoken word recognition (e.g., Blumenfeld & Marian, 2005), and speech production (e.g., Amengual, 2012, 2016). Cognate effects in word recognition have been observed using different experimental paradigms such as lexical decision tasks (e.g., Carrasco-Ortiz et al., 2019; Dijkstra et al., 2010; Van Hell & Dijkstra, 2002), visual-world paradigm eye-tracking tasks (e.g., Blumenfeld & Marian, 2005), reading tasks (e.g., Aguinaga Echeverría, 2017), translation tasks (e.g., Tecedor, 2010), word association tasks (e.g., Van Hell & Dijkstra, 2002), and priming tasks with electrophysiological data (e.g., Comesaña et al., 2012), which shows the robustness of the cognate facilitation effect.

Cross-linguistic orthographic overlap has a consistent role in the facilitation effect of cognates. Otwinowska and Szweczyk (2019) investigated the acquisition of cognates and noncognates in Polish learners of English. Participants completed a backward translation task and a rating task in which they rated the confidence of their translations. The task included cognates, noncognates, and interlingual homographs. Results showed that the cognate facilitation effect was modulated by the degree of orthographic similarity between translation pairs. L2 cognate words were learned faster as orthographic overlap increased. Similarly, orthographic similarity facilitation effects were found in Aguinaga Echeverría (2017). Aguinaga Echeverría (2017) investigated how cognate similarity affected cognate recognition in L1 English–L2 Spanish learners. Participants completed a reading task in Spanish and a translation task that contained cognates with varying degrees of orthographic similarity. Results showed facilitation effects of cognates over noncognates, and cognates with higher orthographic overlap yielded more correct responses. Higher orthographic overlap facilitated recognition of cognates in these studies.

In addition to cross-linguistic orthographic overlap, the degree of phonological overlap has also been found to influence word recognition. However, studies investigating the role of phonological overlap in word recognition have measured phonological similarity in several ways. Some studies established phonological overlap by calculating the number of similar phonetic features (Marian & Spivey, 2003a, 2003b) whereas others calculated overlap based on the number of syllables, position of stressed syllable, vowel quality, and phonological context (Comesaña et al., 2012). These measures may be more appropriate for language pairs that have similar phonetic inventories than for language pairs that differ phonologically in a significant way.

Studies have also used subjective similarity rating tasks to determine the degree of overlap among cognates and noncognates (e.g., Comesaña et al., 2012; Dijkstra et al., 1999, 2010). Subjective ratings may tap directly into the actual similarity that listeners perceive. Other studies have used objective measures of string similarity to calculate both orthographic and phonological overlap of cognate and noncognate pairs such as the Levenshtein Distance (LD) (Carrasco-Ortiz et al., 2019), the Normalized Levenshtein Distance (NLD) (Schepens et al., 2012, 2013), ALINE (Frances et al., 2021; Kondrak, 2000), COGIT (Kondrak, 2001), among others. Relevant to the present study, the LD and the NLD establish form similarity differently, especially for phonological overlap. Whereas the LD considers the number of operations needed to turn one string into another, the NLD considers the distance between sounds based on articulatory features, which

makes the NLD a more appropriate measure of phonological overlap than the LD. Importantly, studies have reported that perceptual similarity ratings of bilinguals and monolinguals commonly correlate with objective measures of overlap (Burt et al., 2017; Comesaña et al., 2012; Dijkstra et al., 2010; Gooskens & Heeringa, 2004; Sanders & Chin, 2009). Notably, evidence of how the NLD correlates with auditory ratings of cross-linguistic phonological similarity is scarce, especially for Spanish–English items.

The effects of cross-linguistic phonological overlap on word recognition may depend on the level of language proficiency. Blumenfeld and Marian (2005) investigated the effects of phonological overlap and language proficiency in bilingual auditory recognition of cognates. English–German and German–English bilinguals completed an auditory identification eye-tracking task in which they identified English cognates and non-cognates in picture displays that also contained German competitor words. The German competitor words shared phonological onsets with the English items. Results showed that German dominant participants co-activated German competitor words when processing both cognate and non-cognate English items; however, English dominant participants exhibited co-activation of German competitor words only when processing English cognates. Both proficiency and phonological overlap modulated the extent of cross-linguistic co-activation of lexical items. Proficiency effects in cognate recognition have also been found in multilingual subjects. Van Hell and Dijkstra (2002) examined the influence of foreign language knowledge on native language performance. Dutch–English–French trilinguals (low and high proficiency) completed a word association task and a lexical decision task in Dutch (L1) in which they encountered L1–L2 cognates, L1–L3 cognates, and L1 non-cognates. Results showed that increased proficiency caused faster responses for cognates over non-cognates. Low proficiency participants did not exhibit cognate advantages, suggesting that a minimum level of proficiency is required to exploit and benefit from cross-linguistic co-activation during L1 processing specifically. However, the role of proficiency in cognate recognition is still unclear since there are also studies showing stronger cognate facilitation effects with lower L2 proficiency (e.g., Bultena et al., 2014; Pivneva et al., 2014; Winther et al., 2021).

Phonological effects of cognates have been observed mostly when processing cognates in the L2 than in the L1. Previous studies have reported an inverse phonological effect in which cognates with greater phonological overlap facilitated processing in the L2 through co-activation of L1 cognate items, but the same effect was not found when processing the L1 cognate items (Carrasco-Ortiz et al., 2019; Muntendam et al., 2022). This processing asymmetry is also observed in priming studies in bilinguals in which cognate effects were found only for L2 cognates primed by L1 cognate translations but not vice versa (Midgley et al., 2009). L1–L2 processing and activation asymmetries are accounted for in different models of lexical activation such as the Revised Hierarchical Model (RHM) (Kroll & Stewart, 1994). For example, in the RHM, the L1 lexicon mediates the connection between the L2 lexicon and the conceptual level, which implies that the L1 influences L2 processing but not vice versa. Hence, cognate facilitation effects are more commonly observed during L2 processing due to strong co-activation of the L1. However, the RHM may not account for cognate facilitation effects in L1 processing.

Since both orthographic and phonological overlap modulate word recognition, studies have explored the interplay of

orthography and phonology in the recognition of cognates. Dijkstra et al. (2010) tested Dutch–English bilinguals using a lexical decision task and a language decision task in which participants read Dutch–English cognates and non-cognates that differed in their level of orthographic and phonological overlap. Results indicated that in the lexical decision task higher orthographic overlap caused faster responses, and increased phonological overlap facilitated recognition of identical cognates only. The language decision task revealed a cognate inhibition effect that increased with orthographic overlap, opposite to the lexical decision task. Orthographic and phonological overlap influenced visual word recognition, but the effects varied depending on task demands. The cognate facilitation effect and the influence of orthographic overlap can be reversed depending on the experimental conditions. Similarly, in Carrasco-Ortiz et al. (2019), English learners of Spanish and Spanish heritage speakers completed two lexical decision tasks (one in Spanish and one in English) in which they read cognates, non-cognates, and pseudo-words that differed in the degree of orthographic and phonological overlap. Results showed that orthographic similarity facilitated word recognition in both languages for both groups, but phonological similarity only facilitated recognition of Spanish words. Phonological and orthographic similarity influenced word recognition in different ways independently of language dominance.

Facilitation effects of form overlap can also depend on the composition of the stimulus list included in experimental tasks. Comesaña et al. (2015) explored the effects of cross-linguistic similarity and stimulus list composition in word recognition in Catalan–Spanish bilinguals. Participants completed a lexical decision task that included identical and non-identical cognates in the stimulus list and a lexical decision task without identical cognates. Importantly, cognate advantages were observed only in the task that included identical cognates in the stimulus list, which highlights the influence of linguistic context in bilingual lexical activation and recognition. Similar to the findings in Dijkstra et al. (2010) and Carrasco-Ortiz et al. (2019), results of Comesaña et al. (2015) showed facilitation effects for orthographic overlap and inhibition effects for phonological overlap. The results point to representational differences between identical cognates and non-identical cognates as well as differential roles of orthographic and phonological overlap. The present study includes identical cognates in the stimulus list to capture facilitation effects in auditory recognition of cognates.

The differences in the effects of phonological and orthographic overlap may depend on the modality of the task, visual or auditory. Frances et al. (2021) investigated the effects of form similarity in L2 word recognition across modalities. Late Spanish–English bilinguals completed a visual lexical decision task and an auditory lexical decision task in English. Both tasks included words with high and low phonological and orthographic similarity, fully crossed, as well as identical cognates. Results indicated that form overlap effects vary depending on the modality of the task. Higher orthographic overlap facilitated visual recognition, and higher phonological overlap facilitated auditory recognition. Interestingly, inhibitory effects were observed across modalities and types of similarity. Higher orthographic overlap hindered auditory recognition and higher phonological overlap hindered visual recognition. The results suggest a need for a separation between types of similarity in cognate studies due to the activation of both types of similarity in both modalities. Other studies have also reported facilitation effects of cognates in the auditory

modality in Spanish–English bilinguals at lower levels of L2 proficiency but have failed to detect any influence of phonological overlap (Andras et al., 2022). Interestingly, more recent studies have reported that cognate recognition in the auditory modality may be influenced by other phonological factors such as lexical stress (Muntendam et al., 2022) and speaker accent (Frances et al., 2022). More research is needed to further explore the factors that modulate cognate facilitation effects in auditory word recognition.

Lexical frequency influences cognate processing as well. Higher lexical frequency cognates have obtained faster responses than lower frequency cognates in visual word recognition (Carrasco-Ortiz et al., 2019; Peeters et al., 2013). However, the way lexical frequency interacts with phonological and orthographic overlap is still unclear. In the visual lexical decision task of Carrasco-Ortiz et al. (2019), lexical items with higher frequency obtained overall faster response times. However, some items with high phonological and orthographic overlap obtained slower response times, which was attributed to lexical frequency. In addition, in Peeters et al. (2013), French–English bilinguals completed a lexical decision task in English. Electrophysiological data and response time data were measured. The results showed facilitation effects of cognates, and reaction times decreased with higher lexical frequency in English and French. This effect showed in the N400 amplitude as well. Importantly, all the cognate items used in Peeters et al. (2013) were identical cognates, so there were not different degrees of overlap involved. It seems that frequency effects are clear and robust with cognates that have complete overlap. When different degrees of overlap are present, frequency effects may diminish the cognate facilitation effect (Carrasco-Ortiz et al., 2019).

In summary, a variety of studies on cognate recognition and processing showed that the degree of orthographic overlap, phonological overlap, lexical frequency, task demands, and language proficiency modulate the facilitation effects of cognates. Orthographic and phonological overlap may have different contributions on cognate effects, but lexical frequency may diminish the effects of cross-linguistic form overlap. It is important to note that most of the studies available address cognate recognition and processing in the visual modality with only a few exceptions addressing the auditory modality. It remains unclear how phonological and orthographic overlap affect auditory word recognition; and it is not clear how proficiency modulates the effects of overlap in the recognition of cognates since many studies did not measure proficiency with objective measures but used self-reporting instead. More research is needed in order to assess the role of form overlap, frequency, and proficiency in recognition of cognates in the auditory modality.

### *Cognates in models of word recognition*

The influence of lexical frequency and orthographic/phonological similarity are accounted for in major models of word recognition. For example, the Bilingual Interactive Activation Plus model (BIA+; Dijkstra & Van Heuven, 2002) considers that activation of lexical items is modulated by the degree of orthographic, phonological, and semantic overlap within and across languages. Frequency of use and neighborhood density also influence lexical processing. However, BIA+ is a model designed to explain visual word recognition and it focuses mainly on orthographic representations. Visual word recognition and spoken word recognition involve different processes due to the nature of spoken and written language.

Although BIA+ acknowledges cognate facilitation effects, the model does not implement the representation of cognates in its design. In the case of cognates, the assumptions of BIA+ predict that higher phonological overlap, orthographic overlap, and lexical frequency result in faster word recognition. Nonetheless, studies suggest that the interactions between these factors are far more complex and that phonological and orthographic cognate similarity may independently cause differential effects (e.g., Carrasco-Ortiz et al., 2019; Comesaña et al., 2015; Dijkstra et al., 1999; Frances et al., 2021).

The assumptions of the BIA+ were further developed into the Multilink model (Dijkstra et al., 2018), which is a localist-connectionist model that simulates the recognition and production of cognates and non-cognates. Multilink assumes non-selective access and parallel activation of word form neighbors across languages. It considers effects of lexical similarity, cognate status, L2 proficiency, word length, lexical frequency, task demands, and translation direction. Regarding lexical similarity, Multilink predicts that words are activated depending on the degree of orthographic similarity to the input word and their subjective frequency of usage. To account for orthographic similarity, Multilink uses the normalized Levenshtein distance to calculate the levels of orthographic overlap across items, and it predicts that lower Levenshtein distance (less orthographic overlap) results in longer recognition times. Word forms in Multilink have a frequency-dependent resting level of activation, which may vary depending on the level of L2 proficiency of individuals. Lexical items in the lexicon of high proficiency bilinguals have a higher resting level of activation than those in the lexicon of low proficiency bilinguals. The frequency-dependent resting level of activation modulates word recognition, with words with higher levels being more strongly activated. Importantly, although Multilink includes cognate effects and predicts the experimental data of several studies employing various tasks, it was designed to account for visual word recognition and production. Similar to the BIA+, Multilink focuses on orthographic representations and visual input, but the model could be adapted to account for auditory word recognition as well.

There are other models of word recognition that may also provide accurate predictions for auditory processing of cognates. The Bilingual Language Interaction Network for Comprehension of Speech (BLINCS: Shook & Marian, 2013) is a computational model that assumes different interconnected levels of processing created by dynamic self-organizing maps (unsupervised learning algorithm). Speech comprehension is influenced by audio-visual integration and cross-linguistic interaction. BLINCS assumes four levels of representation: phonological, phono-lexical, ortho-lexical, and semantic. BLINCS considers that a high degree of cross-linguistic interactions occur (co-activation and competition), which are influenced by lexical frequency and neighborhood size. Shared phonology, semantics, and orthography cause increased activation of candidates within and across languages and lexical selection occurs through the matching of the candidates and the input on the different levels. Regarding cognates, BLINCS assumes that cognates have increased, and stronger activation compared to non-cognates due to orthographic, phonological, and semantic overlap. In this model, cognates are represented near each other in the boundaries of language spaces. This model also considers that proficiency modulates competition and activation. For the present study, BLINCS predicts that cognates will have facilitation effects compared to non-cognates due to cross-language activation and that the level of overlap

and L2 proficiency will modulate these effects. However, similar to the BIA+, BLINCS does not account for the differential contributions of orthographic and phonological overlap previously observed in cognate studies (e.g., Carrasco-Ortiz et al., 2019).

Another available model that gives specific predictions regarding phonological overlap is the Bilingual Interactive Model of Lexical Access (BIMOLA: Léwy & Grosjean, 2008). In BIMOLA, there are two independent lexicons that are interconnected with three levels of representation: a feature level (shared between languages), a phoneme level (language independent but interconnected), and a word level (language independent but interconnected). Cross-linguistic competition happens at the feature level whereas within-language competition happens during the phoneme and word level. In BIMOLA, the language mode of a bilingual listener controls cross-linguistic interactions. This model assumes that phonological similarity determines activation of candidates and that the phonemic repertoires of both languages are interconnected. Phonemes that share only some phonological characteristics (e.g., /b/ and /d/) will be more strongly activated than phonemes that have very different characteristics (e.g., /b/ and /s/). BIMOLA does not implement proficiency effects, but since it has a bilingual feature level with a combined inventory of features from both languages, one could assume that a significant level of proficiency in both languages is needed for the bilingual feature inventory to be established. Frequency effects and the degree of activation are taken into consideration assuming that higher frequency causes stronger activation. BIMOLA does not give specific predictions regarding cognates, but it supports stronger co-activation of overlapping lexical items.

A model that can explain the asymmetry of cognate effects in the L1 and the L2 is the Revised Hierarchical Model (RHM: Kroll & Stewart, 1994). RHM considers that there is a strong connection between the L1 lexicon and the conceptual level, but an indirect connection between the L2 lexicon and the conceptual level. RHM implies that L2 processing is mediated by the L1, but at high levels of proficiency, L2 connections to concepts become stronger. The L1-mediated system of RHM predicts that cognate effects are stronger during L2 processing than during L1 processing, a claim that the present study also tests. Some previous studies have found cognate effects in both languages (e.g., Carrasco-Ortiz et al., 2019; Midgley et al., 2011), but others have found cognate effects only in the L1 (e.g., Caramazza & Brones, 1979; Gerard & Scarborough, 1989). However, some assumptions of the RHM could have different implications for cognate recognition. For example, RHM claims that as proficiency increases L2 lexical items strengthen their direct access to the conceptual level without mediation of the L1, which could imply that as proficiency increases lexical access becomes selective and cognate facilitation effects become weaker. In addition, RHM assumes separate lexicons for the L1 and the L2, which could be problematic for the case of cognates because cognate facilitation effects show strong cross-linguistic interactions. RHM does not explicitly account for cognate representation and the role of orthographic/phonological overlap and lexical frequency.

### The present study

The present study investigates cross-linguistic effects of cognate similarity in auditory recognition in L1 and L2 speakers of Spanish and English. First, this study examines whether objective measures of similarity (LD and NLD) correlate with the subjective similarity ratings of listeners. Previous studies found a positive

correlation between the LD and perceptual distance in Norwegian (Gooskens & Heeringa, 2004) and English (Burt et al., 2017; Sanders & Chin, 2009), and a positive correlation between objective and subjective measures of overlap in Portuguese–English bilinguals (Comesaña et al., 2012) and Dutch–English bilinguals (e.g., Dijkstra et al., 2010; Schepens et al., 2013), but evidence from English–Spanish is still scarce. The present study expects to find a correlation between the LD/NLD and subjective similarity ratings.

Second, this study examines whether cognates are recognized faster and more accurately than non-cognates in L1 and L2 auditory recognition. Evidence of cognate effects comes mostly from visual tasks, and evidence in the auditory modality is scarce. Studies showed that cognate effects are present during L1 and L2 word recognition (e.g., Midgley et al., 2011; Van Hell & Dijkstra, 2002), but others have failed to find facilitation effects in the L1 (e.g., Caramazza & Brones, 1979; Gerard & Scarborough, 1989). The present study expects to find cognate facilitation effects, especially in L2 recognition.

Third, this study tests the influence of orthographic and phonological overlap on auditory recognition of cognates. Models like BIA+ (Dijkstra & Van Heuven, 2002), Multilink (Dijkstra et al., 2018), BLINCS (Shook & Marian, 2013), and BIMOLA (Léwy & Grosjean, 2008) agree that higher cross-linguistic overlap results in stronger co-activation and faster recognition. Some studies have found evidence to support this claim (e.g., Dijkstra et al., 2010; Marian & Spivey, 2003a, 2003b), but others have found that orthographic and phonological overlap may contribute to word recognition in different ways (e.g., Carrasco-Ortiz et al., 2019; Comesaña et al., 2015; Dijkstra et al., 1999; Frances et al., 2021; Schwartz et al., 2007). The present study anticipates orthographic and phonological overlap to modulate auditory recognition of cognates. In auditory recognition, phonological overlap may cause facilitation effects whereas orthographic overlap may cause inhibitory effects or no effect at all, as in previous studies (e.g., Frances et al., 2021).

Finally, this study also evaluates the role of L2 proficiency in auditory recognition of cognates. BLINCS and Multilink predict that stronger cross-linguistic co-activation happens at higher levels of L2 proficiency. Some studies reported cognate facilitation effects at higher levels of L2 proficiency (e.g., Blumenfeld & Marian, 2005; Dijkstra et al., 2010), but others showed facilitation effects at lower levels of L2 proficiency (e.g., Bultena et al., 2014; Pivneva et al., 2014). The present study expects cognate facilitation effects in the L1 to emerge at high levels of L2 proficiency. When processing cognates in the L2, facilitation effects may be more robust with lower L2 proficiency.

## Methodology

### Participants

The sample consisted of 125 participants: 62 Spanish-speaking L2 learners of English (henceforth L1 Spanish subjects) and 63 English-speaking L2 learners of Spanish (henceforth L1 English subjects). An a priori power analysis showed that a total sample of at least 108 participants with two equal sized groups of  $n = 54$  was required to achieve a power of  $\geq .80$  and detect a fixed effect with size .06. Participants were recruited from a university in the southeastern US and through the Prolific.ac online platform. The L1 Spanish subjects were born and raised in Mexico and had no intensive exposure to English before puberty. The

L1 English subjects were born and raised in the US and had no intensive exposure to Spanish before puberty. All participants did not have intensive exposure to languages other than Spanish and English. At the time of testing, participants were living in their country of origin. A background questionnaire gathered information about participants' linguistic profiles and was delivered in the participants' L1 as a survey.

Both groups included speakers with different levels of L2 proficiency, which was measured using the LexTALE test (Lemhöfer & Broersma, 2012) and the LexTALE-ESP test (Izura et al., 2014; Lemhöfer & Broersma, 2012). A description of these tests is provided in the Materials section below. In the LexTALE test in English, scores can range from 0 to 100. L1 English subjects had a mean raw score of 93.5 ( $SD = 6.08$ ), and L1 Spanish subjects had a mean score of 75.1 ( $SD = 9.95$ ). In the LexTALE-ESP test in Spanish, scores can range from -20 to 60. L1 English subjects had a mean raw score of 9.15 ( $SD = 13.4$ ), and L1 Spanish subjects had a mean score of 48.4 ( $SD = 6.66$ ). The L1 Spanish subjects had higher L2 proficiency than the L1 English subjects, but both samples included a wide range of scores. Figure 1 shows the raw scores of both proficiency tests for both groups.

## Materials and procedure

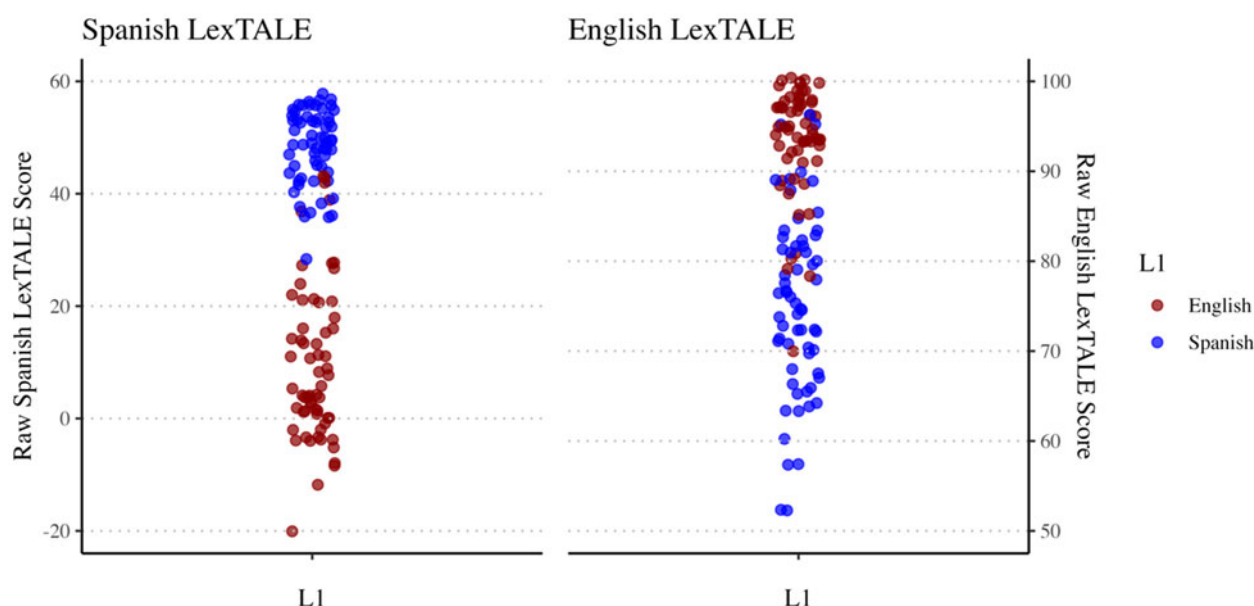
### Proficiency tests

Participants completed two tests that assess proficiency via vocabulary size. The LexTALE test (Lemhöfer & Broersma, 2012) provided a measure of participants' English proficiency by assessing vocabulary size through a lexical decision task. Participants saw 40 real words and 20 pseudowords and indicated whether the items are real words in the English language. Scores can range from 0 to 100, and incorrect responses are penalized. The LexTALE-ESP test (Izura et al., 2014) is a modified version of the LexTALE that measures Spanish vocabulary size. Participants saw 60 real words and 30 pseudowords and indicated whether the items are existing words in Spanish. Scores can range from -20 to 60, and incorrect responses are penalized.

### Auditory lexical decision tasks

Participants completed two Auditory Lexical Decision Tasks (ALDT), in Spanish and English. Participants listened to cognates, non-cognates, and pseudowords and indicated whether these items are real words. Trials began with the words 'Real' and 'Not Real' displayed on the sides of the screen, and 500 ms after, a word was presented aurally. Participants answered by pressing key '1' (Real) or '0' (Not real). Accuracy and response times (RT) were measured from the offset of the stimulus word. There were five practice trials to familiarize participants with the experimental procedure. Instructions were given in the language of the ALDT.

The stimuli items were 100 cognate pairs and 100 noncognate pairs, which were matched for orthographic length and lexical frequency. Orthographic length and lexical frequency were calculated using the web application NIM (Guasch et al., 2013). Orthographic length in English cognates (mean = 5.13,  $SD = 1.03$ ) and noncognates (mean = 4.96,  $SD = .91$ ) was not statistically different:  $t(194.93) = 1.24$ ,  $p = .218$ . Similarly, orthographic length in Spanish cognates (mean = 5.57,  $SD = 1.10$ ) and noncognates (mean = 5.46,  $SD = 1.39$ ) was not statistically different:  $t(188.39) = .62$ ,  $p = .536$ . Since NIM uses different corpora for lexical frequency in Spanish and English, frequency values were converted to the Zipf scale, which is a standardized scale of



**Figure 1.** LexTALE raw scores. The left panel plots raw LexTALE-ESP scores (Spanish vocabulary test). The right panel plots raw LexTALE scores (English vocabulary test).

word frequency per million logged to the base of 10 + 3 (Van Heuven et al., 2014). Lexical frequency of English cognates (mean = 1.70,  $SD = .46$ ) and noncognates (mean = 1.64,  $SD = .56$ ) was not statistically different:  $t(190.39) = .83$ ,  $p = .408$ , and frequency of Spanish cognates (mean = 1.61,  $SD = .40$ ) and noncognates (mean = 1.48,  $SD = .60$ ) was not statistically different:  $t(173.4) = 1.8$ ,  $p = .073$ .

Cognates and non-cognates had varying degrees of orthographic and phonological overlap, which was calculated using the Levenshtein Distance (LD) (Yarkoni et al., 2008) and the Normalized Levenshtein Distance (NLD) (Schepens et al., 2012). The LD calculates string similarity based on the number of substitution, insertion, or deletion operations needed to turn one string into the other. For example, for *air* and *aire*, the orthographic LD is 1 due to the insertion/deletion of “e”. A higher LD value corresponds to lower orthographic overlap because a greater number of operations is needed. Previous studies have used the LD to measure phonological overlap as well (e.g., Carrasco-Ortiz et al., 2019; Vitevitch, 2012) by applying it to the phonetic transcription of lexical items. For instance, for *air* / $\epsilon r$ / and *aire* / $\text{aire}$ /, the phonological LD value is 4 because all four segments have to be modified to turn one transcription into the other. Because the LD does not consider differences in the features of phonemes, the present study also used the NLD, which is a normalization of the LD that has values between 0 and 1. Lower

NLD corresponds to lower overlap. For phonological overlap, this study employed the phonological NLD adapted by Schepens et al. (2013). Phonological NLD includes different substitution costs based on the distinctive phonetic feature space as presented in the International Phonetic Alphabet (IPA). Schepens et al. (2013) calculated the Euclidean distance in the respective IPA vowel or consonant space and applied the respective substitution costs for articulatory similarity to the phonetic transcriptions of the lexical items. The phonetic transcriptions of English items were obtained using the web application *toPhonetics*, and the Spanish transcriptions were obtained using the *Transcriptor Castellano* of The Musa Academy.

All the cognates included in the present study had values of more than .5 orthographic NLD and .6 phonological NLD, and the noncognates had values of less than .3. Orthographic NLD for English–Spanish cognates (mean = .78,  $SD = .13$ ) and noncognates (mean = .12,  $SD = .095$ ) were statistically different:  $t(179.68) = 40.875$ ,  $p < .001$ ; and the phonological NLD of cognates in the present study was on average .81 ( $SD = .08$ ). Orthographic NLD and LD were calculated using the web application NIM (Guasch et al., 2013). Phonological LD was obtained using the VWR package in R (Keuleers, 2013), and phonological NLD was obtained from the *Dataset S1* provided by Schepens et al. (2013). All measures can be found in the Supplementary Materials. Table 1 summarizes the lexical properties of the experimental stimuli.

**Table 1.** Lexical properties of the experimental stimuli (means with standard deviations displayed in parentheses)

| Language | Word type  | Orthographic length | Zipf frequency     | Orthographic NLD  | Phonological NLD  | Orthographic LD     | Phonological LD     |
|----------|------------|---------------------|--------------------|-------------------|-------------------|---------------------|---------------------|
| Spanish  | Cognate    | 5.57 ( $\pm 1.10$ ) | 1.61 ( $\pm .40$ ) | .78 ( $\pm .13$ ) | .81 ( $\pm .08$ ) | 1.25 ( $\pm .82$ )  | 3.87 ( $\pm 1.15$ ) |
|          | Noncognate | 5.46 ( $\pm 1.39$ ) | 1.48 ( $\pm .60$ ) | .12 ( $\pm .09$ ) | .01 ( $\pm .02$ ) | 5.14 ( $\pm 1.03$ ) | 5.3 ( $\pm 1.28$ )  |
| English  | Cognate    | 5.13 ( $\pm 1.03$ ) | 1.70 ( $\pm .46$ ) | .78 ( $\pm .13$ ) | .81 ( $\pm .08$ ) | 1.25 ( $\pm .82$ )  | 3.87 ( $\pm 1.15$ ) |
|          | Noncognate | 4.96 ( $\pm .91$ )  | 1.64 ( $\pm .56$ ) | .12 ( $\pm .09$ ) | .01 ( $\pm .01$ ) | 5.14 ( $\pm 1.03$ ) | 5.3 ( $\pm 1.28$ )  |

Experimental items were divided into two presentation lists per language. Each list contained 50 cognates and 50 noncognates. The presentation lists were counterbalanced so that participants do not encounter a Spanish word and its English translation across both ALDTs. A total of 100 pseudo-words were included in each ALDT, a ratio of words to non-words of 1:1 as in previous studies (Comesaña et al., 2015; Dijkstra et al., 2010). The pseudo-words in the ALDTs were orthographically and phonologically legal constructions in Spanish and English.

### Similarity rating task

Participants completed a Similarity Rating Task (SRT) in which they rated the phonological similarity between cognates using a six-point Likert scale (1 = totally different, 6 = totally similar). Participants were instructed to focus strictly on how the words sound and ignore spelling and meaning. All participants listened to the same words used in the ALDTs. Trials began with the rating options displayed on the screen. 500 ms after, a Spanish word was presented aurally followed by its English translation. Participants' ratings were recorded using a computer keyboard.

### Data analysis

The ratings obtained in the SRT were submitted to bivariate correlation analyses with the LD and NLD values previously calculated. All values were converted to standardized scores (z-scores) so that they are in the same scale and have a mean of 0. Correlation analyses calculate a Pearson's correlation coefficient, which provides values between 1 and -1. Values closer to 0 indicate weaker/lower correlations. The correlation analyses were conducted using the *Stats* package in R (R Core Team, 2022).

Both ALDTs measured response times and accuracy (correct/incorrect responses). Correct responses were automatically coded as "1" and incorrect responses as "0." Following Carrasco-Ortiz et al. (2019), RTs of incorrect responses and those shorter than 100 ms and 2.5 SD above the participant mean were excluded from the analysis. Outliers or influential data points were detected by calculating the Cook's distance measure. RTs and correct responses were analyzed using mixed effects models, including the fixed effects *L1 group* (English, Spanish), *word type* (cognate, noncognate, pseudoword), *phonological overlap* (LD, NLD, similarity ratings), *orthographic overlap* (LD, NLD), and *proficiency*. The models applied treatment contrast coding. The logistic models used *L1 group:English* and *Word type:cognate* as the intercept, and the linear models used *L1 group:English* as the intercept. The random effects structure included by-subject and by-word random intercepts and by-subject and by-word random slopes. Main effects and interactions were evaluated by hierarchically partitioning the variance via nested model comparisons. Relevant interactions were assessed by computing the estimated marginal means of the associated predictors and the comparisons among them. Alpha was set at .05. All analyses were carried out using R (R Core Team, 2022). The package *lme4* (Bates et al., 2015) was used to fit the mixed effects models, and the package *emmeans* (Lenth, 2021) was used for multiple comparisons. The scripts used for the analysis of the data are available at: <https://osf.io/765ytr/>.

## Results

### Similarity rating task

The ratings of L1 Spanish subjects and L1 English subjects were very strongly positively correlated,  $r(98) = .87$ ,  $p < .001$ . Since

both groups rated the phonological similarity of the cognate pairs in a very similar way (see Figure 2), the ratings of both groups were merged into one variable, and a mean rating was calculated for each cognate pair. Mean ratings of phonological similarity were weakly negatively correlated with the phonological LD values,  $r(98) = -.24$ ,  $p = .014$ ; and moderately negatively correlated with the orthographic LD values,  $r(98) = -.40$ ,  $p < .001$ . In addition, mean ratings were moderately positively correlated with the phonological NLD values,  $r(98) = .52$ ,  $p < .001$ ; and moderately positively correlated with the orthographic NLD values,  $r(98) = .50$ ,  $p < .001$ . The similarity ratings were not correlated with the Zipf frequency values,  $r(98) = .086$ ,  $p = .39$ . Figure 3 shows all the relevant correlations.

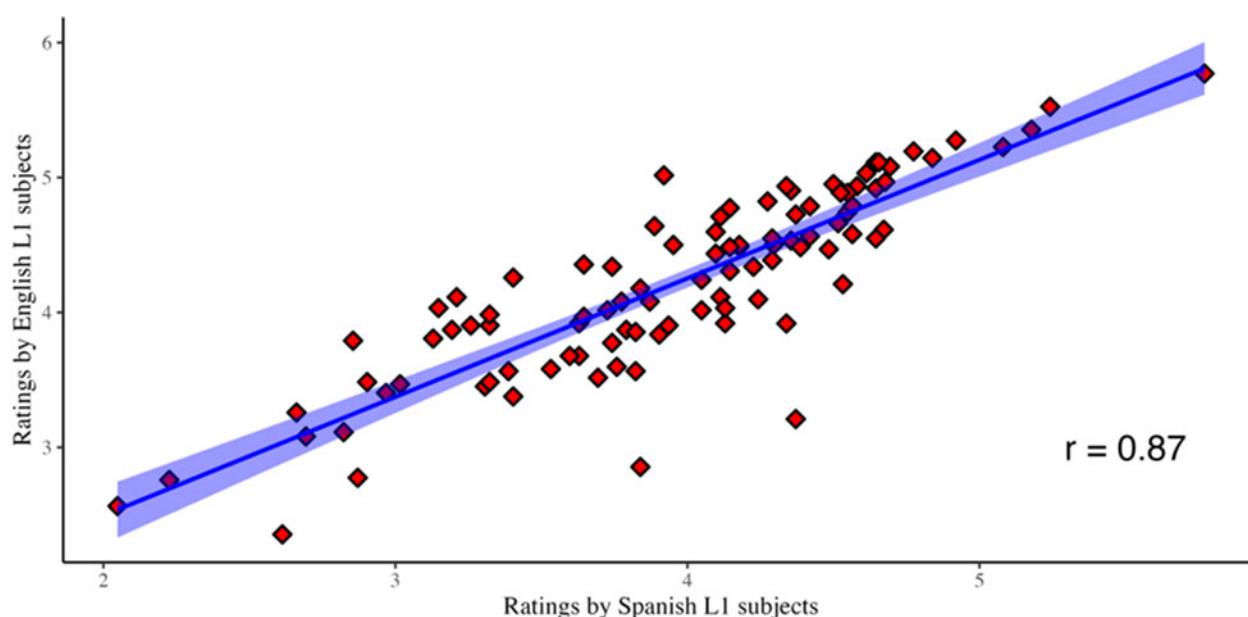
### Auditory lexical decision task in Spanish

Summaries of the output of the models can be found in the Supplementary Materials. When testing cognate effects on the proportion of correct responses, the models revealed a main effect of *L1 group* ( $\chi^2(1) = 85.44$ ,  $p < .001$ ) and *word type* ( $\chi^2(2) = 21.73$ ,  $p < .001$ ). There was also a two-way interaction between *L1 group* and *word type* ( $\chi^2(2) = 12.28$ ,  $p = .002$ ). Pairwise comparisons by *L1* and *word type* revealed significant differences in the proportion of correct responses between real words and pseudowords but not between cognates and noncognates for both groups. However, L1 English subjects were slightly more accurate identifying cognates than noncognates, and L1 Spanish subjects were slightly more accurate identifying noncognates than cognates.

For testing the effect of form overlap on the proportion of correct responses, only cognates were included in the analysis of form overlap, so *word type* was excluded from this model. The best fit model did not include random slopes for orthographic or phonological overlap due to convergence issues. The model revealed a main effect of *group* ( $\chi^2(1) = 33.137$ ,  $p < .001$ ), *phonological NLD* ( $\chi^2(1) = 8.04$ ,  $p = .004$ ), and *similarity ratings* ( $\chi^2(1) = 20.38$ ,  $p < .001$ ) but not of *phonological LD* ( $\chi^2(1) = .498$ ,  $p = .48$ ). The model also yielded a significant two-way interaction between *L1 group* and *phonological NLD* ( $\chi^2(1) = 31.302$ ,  $p < .001$ ) and an interaction between *L1 group* and *similarity ratings* ( $\chi^2(1) = 17.816$ ,  $p < .001$ ). The parameter estimates of the model indicated that a one-unit increase in the similarity ratings yielded a change in the log odds of identifying cognates correctly by .74 for L1 English subjects and by -.62 for L1 Spanish subjects. Similarly, a one-unit increase in phonological NLD yielded a change in the log odds by 2.65 for L1 English subjects and -4.55 for L1 Spanish subjects. The model also revealed a main effect of *Spanish proficiency* (LexTALE-ESP score) on the proportion of correct responses ( $\chi^2(1) = 63.32$ ,  $p < .001$ ). Participants provided more correct responses with higher proficiency in Spanish. There were no effects of orthographic LD or orthographic NLD.

On RTs, the model revealed a main effect of *L1 group* ( $\chi^2(1) = 10.92$ ,  $p < .001$ ) and *word type* ( $\chi^2(2) = 186.36$ ,  $p < .001$ ). There was also a significant two-way interaction between *L1 group* and *word type* ( $\chi^2(2) = 7.07$ ,  $p = .029$ ). Similar to the correct response data, pairwise comparisons by *L1* and *word type* revealed significant differences in RTs between real words and pseudowords but not between cognates and noncognates for both groups.

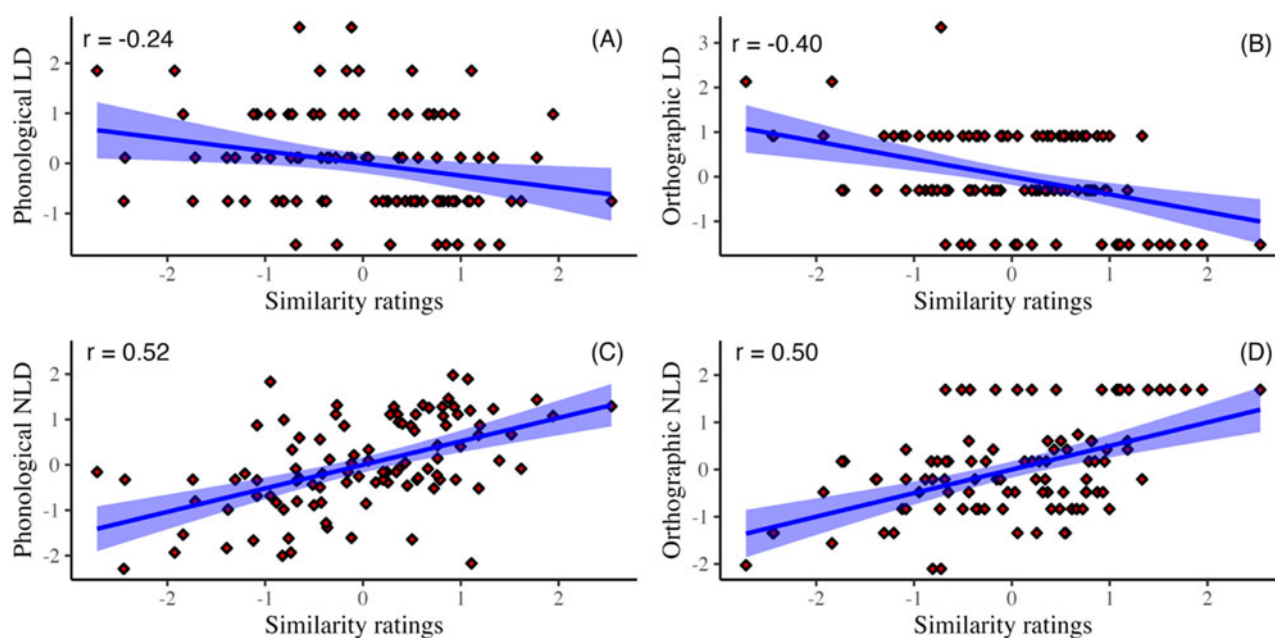
Regarding phonological overlap, the best fit model did not include random slopes for orthographic or phonological overlap due to convergence issues. The model yielded a main effect of



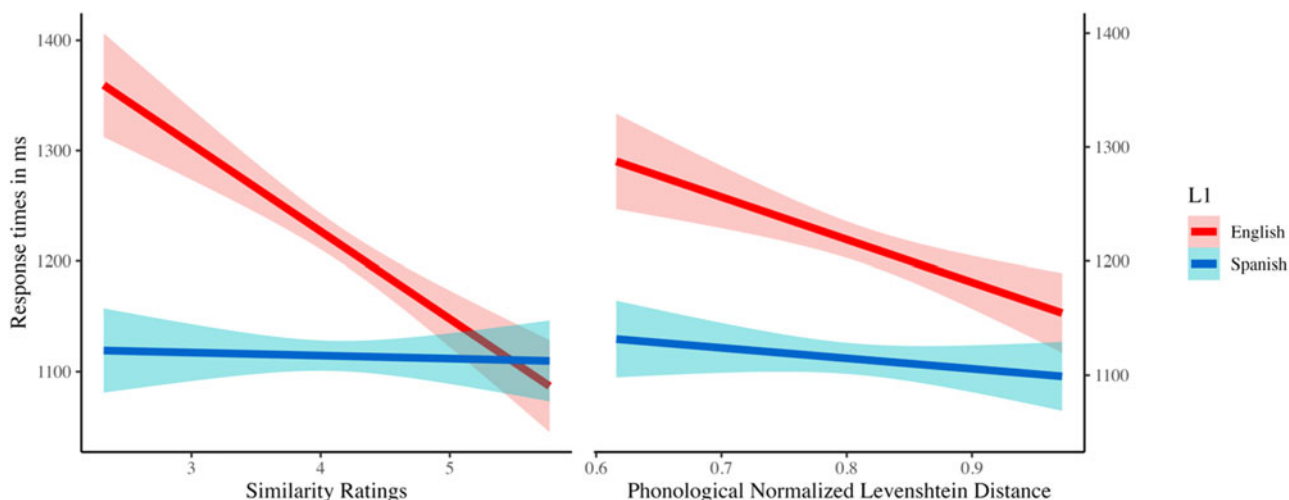
**Figure 2.** Correlation of the similarity ratings of cognate pairs provided by L1 Spanish-L2 English subjects and L1 English-L2 Spanish subjects in the similarity rating task.

*L1 group* ( $\chi^2(1) = 10.84, p < .001$ ) and *similarity ratings* ( $\chi^2(1) = 4.97, p = .025$ ), but not of phonological LD ( $\chi^2(1) = .244, p = .62$ ) or phonological NLD ( $\chi^2(1) = 1.62, p = .203$ ). There was also a significant interaction between *L1 group* and *similarity ratings* ( $\chi^2(1) = 19.1, p < .001$ ). Similarity ratings affected RTs of L1 English subjects more than L1 Spanish subjects (see Figure 4). The parameter estimates of the model showed that on average L1 English subjects responded about 74 ms faster with a one-unit increase in the similarity ratings, and about 30 ms faster with a one-unit increase in phonological NLD. Figure 4 shows RTs for

cognate words for each group based on phonological NLD and similarity ratings. The model also yielded a main effect of *Spanish proficiency* (LexTALE-ESP score) on RTs ( $\chi^2(1) = 14.33, p = .05$ ). Participants in both groups exhibited faster RTs overall with higher Spanish proficiency score. Although L1 English subjects at all levels of proficiency exhibited faster responses with higher phonological overlap, visual examination of the data showed a trend that L1 English subjects with lower Spanish proficiency benefited the most from increasing phonological overlap in cognates, followed by subjects with very high proficiency. There



**Figure 3.** Correlation of phonological similarity ratings and phonological Levenshtein Distance (panel A), orthographic Levenshtein Distance (panel B), phonological Normalized Levenshtein Distance (panel C), and orthographic Normalized Levenshtein Distance (panel D). Values are converted to z-scores.



**Figure 4.** Response times for cognate words per group based on Similarity Ratings (left panel) and phonological NLD (right panel) in the ALDT in Spanish.

were no effects of orthographic *LD* or orthographic *NLD* in RTs. Appendix S1 in the Supplementary Materials presents the summary of the models used to make inferences.

#### *Auditory lexical decision task in English*

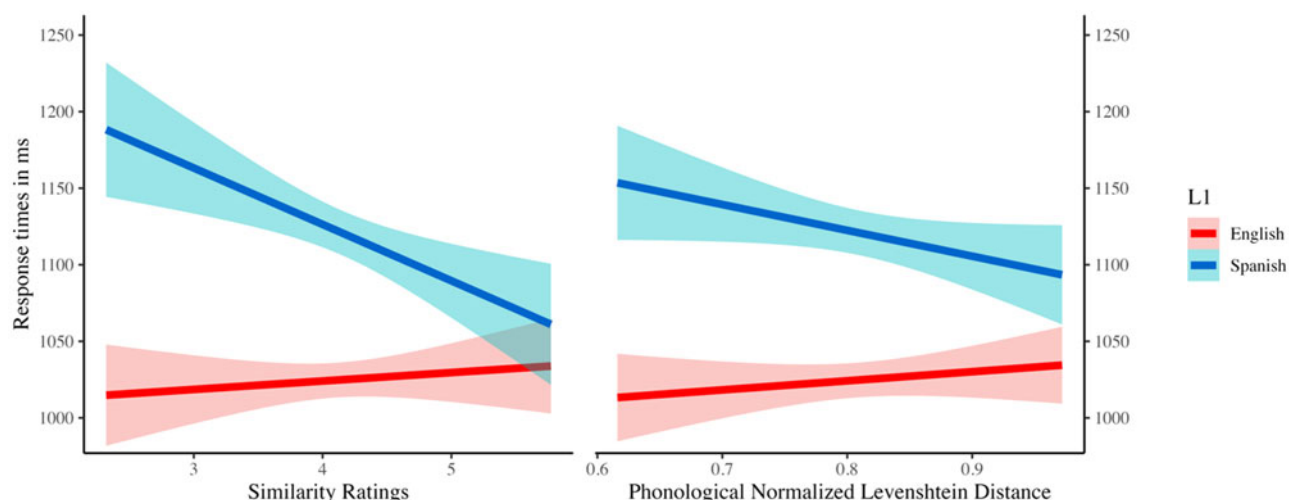
When testing cognate effects on the proportion of correct responses, the models revealed a main effect of *L1 group* ( $\chi^2(1) = 68.64, p < .001$ ) and *word type* ( $\chi^2(1) = 68.67, p < .001$ ) on the proportion of correct responses. There was a significant two-way interaction between *word type* and *L1 group* ( $\chi^2(2) = 45.08, p < .001$ ). Similar to the ALDT in Spanish, pairwise comparisons by *L1* and *word type* revealed significant differences in the proportion of correct responses between real words and pseudowords but not between cognates and noncognates for both groups.

Regarding phonological overlap, the best fit model included by-subject and by-word random slopes for phonological overlap (similarity ratings) as well as by-subject and by-word random intercepts. The model revealed a main effect of *similarity ratings* ( $\chi^2(1) = 8.11, p = .004$ ) but not of phonological *LD* ( $\chi^2(1) = 2.29, p = .129$ ) or phonological *NLD* ( $\chi^2(1) = 1.84, p = .174$ ). There were

no significant interactions. The parameter estimates of the model indicated that a one-unit increase in the similarity ratings yielded a change in the log odds of identifying cognates correctly by .51 for L1 English subjects and by .11 for L1 Spanish subjects. There was also a main effect of *English proficiency* (LexTALE score) on the proportion of correct responses ( $\chi^2(1) = 3.86, p = .049$ ). Participants provided more correct responses with higher proficiency in English. Orthographic *LD* and orthographic *NLD* had no effects on the response data.

In the RT data, the model revealed a main effect of *L1 group* ( $\chi^2(1) = 31, p < .001$ ) and *word type* ( $\chi^2(2) = 225, p < .001$ ). There was also a significant two-way interaction between *L1 group* and *word type* ( $\chi^2(2) = 358, p < .001$ ). Pairwise comparisons by *L1* and *word type* yielded significant differences in RTs between real words and pseudowords but not between cognates and noncognates for both groups.

For the effect of phonological overlap on RTs, the best fit model included by-word random slopes for phonological overlap (similarity ratings) as well as by-subject and by-word random intercepts. The model yielded no main effects of any of the measures of phonological overlap, and there was a significant two-way



**Figure 5.** Response times for cognate words per group based on Similarity Ratings (left panel) and phonological NLD (right panel) in the ALDT in English.

interaction between *L1 group* and *similarity ratings* ( $\chi^2(1) = 5.43$ ,  $p = .019$ ). Similarity ratings affected RTs of L1 Spanish subjects more than L1 English subjects (see Figure 5). The parameter estimates of the model showed that on average L1 Spanish subjects responded about 30 ms faster with a one-unit increase in the similarity ratings. Figure 5 shows RTs for cognate words for each group based on phonological NLD and similarity ratings. The model also yielded a main effect of *English proficiency* (LexTALE score) on RTs ( $\chi^2(1) = 8.65$ ,  $p = .003$ ). Participants in both groups exhibited faster RTs overall with higher English proficiency score. L1 Spanish subjects with lower English proficiency seemed to benefit the most from increasing phonological overlap in cognates. There were no effects of orthographic LD or orthographic NLD in RTs. Appendix S1 in the Supplementary Materials presents the summary of the models used to make inferences.

### Exploratory analysis

The initial results of the present study did not yield any effects of cognate status for either group in either of the tasks. A lack of cognate effects has been reported in several studies (e.g., Comesaña et al., 2012, 2015; Schwartz et al., 2007; VanLangendonck et al., 2020), but the majority of the literature argues in favor of cognate facilitation effects. In order to test whether the lack of cognate effects in the present study was related to any part of the methodological design, a post-hoc exploratory analysis was conducted to examine whether the lack of cognate effects was related to the method used in the present study to distinguish between cognates and noncognates. Currently, there is no consensus regarding how much phonological and orthographic overlap is needed for a word to be considered a cognate and to be able to detect cognate effects. In addition, different studies used different measures to account for form similarity and distinguish between cognates and noncognates. Schepens et al. (2013), for example, used an inclusive threshold of .5 orthographic NLD and .75 phonological NLD while Carrasco-Ortiz et al. (2019) used the LD and Frances et al. (2021) used ALINE distance to account for form similarity. The exploratory analysis manipulated the threshold and the measure to classify cognates and noncognates.

Initially, the study used an inclusive threshold of at least .5 orthographic NLD (mean = .78) and .6 phonological NLD (mean = .81) for cognates. Following Schepens et al. (2013), in the exploratory analysis, the data were recoded with an inclusive threshold of .5 orthographic NLD (mean = .78) and .75 phonological NLD (mean = .84) and reanalyzed to test cognate effects. In the Spanish ALDT, the models revealed a main effect of L1 group ( $\chi^2(1) = 85.44$ ,  $p < .001$ ), word type ( $\chi^2(2) = 25.93$ ,  $p < .001$ ), and a significant interaction between word type and L1 group ( $\chi^2(2) = 42.02$ ,  $p < .001$ ). Pairwise comparisons revealed a facilitation effect of cognates for L1 English participants on the proportion of correct responses ( $p = .006$ ). Cognate effects were not detected in RT data in the Spanish ALDT. In the English ALDT, no cognate effects were detected for either group using this new threshold.

The present study used phonological and orthographic NLD, an objective measure, as the initial measure to select the stimuli words. The exploratory analysis tested whether using perceptual similarity ratings, a subjective measure, to distinguish between cognates and noncognates could better explain the data. The data were recoded and words having at least 4 out of 6 mean similarity rating were coded as cognates. A cognate facilitation effect

was detected on the proportion of correct responses ( $p < .001$ ) and response times ( $p = .006$ ) of L1 English subjects in the Spanish ALDT. In the English ALDT, no cognate effects were detected for cognates with at least 4 out of 6 similarity rating.

The initial analysis protocol of the present study used Pearson's correlation analyses to test the relationship between similarity ratings and objective measures of overlap. During Stage 2 of this registered report, it was noted that Pearson's correlation analyses are not the best method to analyze Likert scale data due to their ordinal nature. For this reason, exploratory Spearman's correlation analyses were conducted to test whether the correlations reported in the present study coincided with this more appropriate method (see Appendix S2 in the Supplementary Materials). The Spearman's correlation coefficients were almost identical to the Pearson's correlation coefficients reported in the present study, which confirms that the relationships between the measures of overlap reported in the current study are accurate under these two methods.

### Discussion

The present study investigated cross-linguistic interactions in auditory recognition of cognates in L1 and L2 Spanish and English. The study examined the role of cross-linguistic phonological overlap in cognate recognition using objective and subjective measures of overlap. The study had four main goals: (1) to test whether objective measures of overlap correlated with subjective similarity ratings, (2) to analyze whether cognate facilitation effects are observed in spoken word recognition, (3) to study the influence of orthographic and phonological overlap on auditory recognition of cognates, and (4) to examine the role of L2 proficiency in cognate recognition. The results of a similarity rating task and two auditory lexical decision tasks showed that there were moderate correlations between objective and subjective measures of overlap, that cognate facilitation effects were not observed in the auditory lexical decision tasks, that auditory recognition of cognates is facilitated by increased phonological overlap, and that higher L2 proficiency correlates with faster and more accurate spoken word recognition.

#### Correlation of subjective and objective measures of cross-linguistic overlap

In the present study, L1 and L2 speakers of Spanish and English rated the degree of similarity of cognate pairs. The ratings provided by L1 Spanish subjects and L1 English subjects showed a strong correlation, which suggested that both groups perceived phonological overlap in a very similar way regardless of their L1/L2 status and their level of proficiency. The overall ratings provided by both groups of participants had a weak correlation with phonological LD, a moderate correlation with orthographic LD, a moderate correlation with phonological NLD, and a moderate correlation with orthographic NLD. Phonological NLD was the measure that best correlated with the perceptual similarity ratings provided by participants in the similarity rating task.

Importantly, the LD and the NLD establish form similarity differently, especially for phonological overlap. Whereas the LD focuses mainly on string similarity, the NLD considers the distance between sounds based on different articulatory features and applies substitution costs for distinctive phonetic features, which makes the NLD a more appropriate measure of phonological overlap. This difference is clearly reflected in how these

two measures correlated with perceptual ratings. However, while studies comparing objective and subjective measures found a strong correlation between the LD and perceptual similarity of monolingual listeners in Norwegian (Gooskens & Heeringa, 2004) and English (Burt et al., 2017; Sanders & Chin, 2009), and a strong correlation between objective and subjective measures in Portuguese–English bilinguals (Comesaña et al., 2012) and Dutch–English bilinguals (e.g., Dijkstra et al., 2010; Schepens et al., 2013), the present study found only a weak correlation of perceptual ratings with phonological LD and a moderate correlation with phonological NLD, suggesting that these measures of overlap do not completely align with each other but are closely related.

### *Cognate facilitation effects in spoken word recognition*

Facilitation effects of cognates over noncognates were not initially detected in either task for either group in the present study. In the Spanish task, there was a trend for L1 English subjects with intermediate Spanish proficiency to recognize cognates slightly faster than noncognates; and in the English task, L1 Spanish subjects with high English proficiency recognized cognates slightly faster than noncognates. However, there was no significant effect of the cognate status in either task, which coincides with previous studies reporting a lack of facilitation effects of cognates (e.g., Comesaña et al., 2012, 2015; Schwartz et al., 2007; VanLangendonck et al., 2020).

Although the lack of facilitation effects of cognates in the present study is supported by several previous studies, the results of the present study call for a reassessment of how cognates and noncognates are classified across different studies. There is no consensus regarding how much orthographic and phonological overlap a cognate must have to cause facilitation effects, and studies differ significantly in how they operationalize cognates. The present study initially regarded as cognates all the translation pairs that had more than .5 orthographic NLD and more than .6 phonological NLD. Other studies have used a different threshold and measure for the distinction of cognates and noncognates. For example, Schepens et al. (2013) used an inclusive threshold of .5 orthographic NLD and .75 phonological NLD, Carrasco-Ortiz et al. (2019) used the LD to account for orthographic and phonological similarity, and Frances et al. (2021) used ALINE distance.

It is possible that among other factors, the presence/absence of cognate effects across studies is related to how cognates are methodologically established in the first place. In the exploratory analysis conducted in the present study, cognate facilitation effects emerged when changing the inclusive threshold from .6 to .75 phonological NLD, following Schepens et al. (2013). These results show that to be able to detect cognate effects in a ALDT, cognates should have at least .75 phonological NLD. Notably, a subjective measure seems to predict more accurately the effects of cognates in spoken word recognition. In the exploratory analysis, when the present study used similarity ratings instead of phonological NLD as the measure to distinguish between cognates and noncognates, with a threshold of at least 4 out of 6, cognate effects emerged in both correct responses and response times. The results show that subjective measures of phonological similarity can predict cognate effects successfully and that slight differences in how cognates and noncognates are established in linguistic studies have a great impact on our ability to detect cognate effects.

### *Effects of cross-linguistic overlap and L2 proficiency on cognate recognition*

The results of the present study revealed that higher phonological overlap, measured objectively and subjectively, resulted in faster and more accurate auditory recognition of cognate words, and that orthographic overlap did not have an effect in either of the tasks. These results support studies reporting evidence of an interaction between the type of form overlap and the modality of the tasks (e.g., Frances et al., 2021); with orthographic overlap facilitating visual word recognition but not auditory word recognition, and phonological overlap facilitating auditory word recognition but not visual word recognition. However, in Frances et al. (2021) similarity across modalities hindered recognition, but in the present study increasing orthographic similarity was not associated with less successful auditory recognition. Importantly, phonological overlap showed a bigger effect in L2 recognition than in L1 recognition, which supports the lexical processing asymmetry represented in the Revised Hierarchical Model (Kroll & Stewart, 1994). Higher phonological overlap facilitated recognition for L2 learners at all levels of proficiency, and learners with lower L2 proficiency in both languages seemed to benefit the most from increased phonological similarity. The results support models of word recognition such as BIA+ (Dijkstra & Van Heuven, 2002), Multilink (Dijkstra et al., 2018), BLINCS (Shook & Marian, 2013), and BIMOLA (Léwy & Grosjean, 2008) which suggest that higher cross-linguistic overlap facilitates lexical activation and recognition. The study also coincides with previous research showing that lexical recognition becomes easier with higher cross-linguistic overlap (e.g., Dijkstra et al., 2010; Marian & Spivey, 2003a, 2003b; Carrasco-Ortiz et al., 2019) and studies that found that orthographic and phonological overlap affect word recognition differently across task modalities (e.g., Carrasco-Ortiz et al., 2019; Frances et al., 2021; Schwartz et al., 2007). Most models currently do not account for modality dependent effects in word recognition which is necessary to understand the interactions between the types of cross-linguistic overlap and the modality in which the input is presented.

The different measures of phonological overlap used in this study differed in how accurate they were at predicting the pattern of lexical decision responses. Phonological LD did not successfully predict response accuracy or response times in either of the tasks. LD is a measure of string similarity that is not the most appropriate for measuring phonological overlap, especially when studying auditory recognition and processing. Previous studies have used LD to measure phonological overlap (e.g., Carrasco-Ortiz et al., 2019; Vitevitch, 2012) and have been able to successfully detect effects, but these studies used visual tasks only. On the other hand, the NLD and the self-rated similarity values were more accurate measures. The NLD successfully predicted response accuracy and response times in the Spanish lexical decision task, but it did not have a significant effect in the English task. The similarity ratings, however, had significant effects in response accuracy and response times in both tasks. Based on these results, the present study showed that self-rated phonological similarity was the most successful measure of cross-linguistic phonological overlap in comparison to the objective measures. It is possible that objective measures take into account aspects that are not salient and relevant at the perceptual level, and they may overestimate the influence of certain phonetic features. Importantly, the present study calculated mean similarity ratings per word combining the ratings of all participants considering the strong correlation

between ratings across both L1 groups. Further research is needed to examine L1 dependent effects and individual differences in how participants rate phonological similarity and how it relates to individual patterns of auditory recognition. Finally, higher proficiency was associated with faster and more accurate auditory recognition. When examining the interplay of proficiency and phonological overlap, there was a trend that showed stronger facilitation effects of phonological overlap at lower levels of proficiency. Future research should dive deeper into the interplay of proficiency and phonological overlap in word recognition.

## Conclusion

The present study examined the influence of cross-linguistic overlap in spoken word recognition by comparing different measures of overlap. The study provided evidence that objective and subjective measures of overlap correlate moderately with each other and that L1 and L2 spoken word recognition is facilitated by increased phonological overlap. Among all the measures of overlap employed, perceived phonological similarity was found to be the most accurate predictor of recognition patterns. Phonological overlap showed larger facilitation effects in L2 recognition than L1 recognition. Cognate effects in auditory word recognition were detected using both subjective and objective measures to distinguish between cognates and noncognates, but these effects were limited only to L1 English subjects processing L2 Spanish. The findings are in line with models of lexical processing that suggest that co-activation of lexical items across languages is modulated by the degree of similarity and previous studies reporting modality dependent effects of cross-linguistic form overlap in lexical recognition.

**Supplementary materials and availability of data.** The data that support the findings of this study and other supplementary materials are openly available in Open Science Framework, which can be accessed at: <https://osf.io/765yr/>

Supplementary material can be found online at <https://doi.org/10.1017/S1366728924000270>

## References

- Aguinaga Echeverría, S. (2017). All cognates are not created equal: variation in cognate recognition and applications for second language acquisition. *Revista Española de Lingüística Aplicada*, 16(1), 23–42.
- Amengual, M. (2012). Interlingual influence in bilingual speech: Cognate status effect in a continuum of bilingualism. *Bilingualism: Language and Cognition*, 15(3), 517–530.
- 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. *Frontiers in Psychology*, 7 (617).
- Andras, F., Rivera, M., Bajo, T., Dussias, P. E., & Paolieri, D. (2022). Cognate facilitation effect during auditory comprehension of a second language: A visual world eye-tracking study. *International Journal of Bilingualism*, 1–21.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48.
- Blumenfeld, H., & Marian, V. (2005). Covert Bilingual Language Activation through Cognate Word Processing: An Eye-tracking Study. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 27.
- Bultena, S., Dijkstra, T., & Van Hell, J. G. (2014). Cognate effects in sentence context depend on word class, L2 proficiency, and task. *The Quarterly Journal of Experimental Psychology*, 67(6), 1214–1241.
- Burt, J. S., McFarlane, K. A., Kelly, S. J., Humphreys, M. S., Weatherall, K., & Burrell, R. G. (2017). Brand name confusion: Subjective and objective measures of orthographic similarity. *Journal of Experimental Psychology: Applied*, 23(3), 320–335.
- Caramazza, A., & Brones, I. (1979). Lexical access in bilinguals. *Bulletin of the Psychonomic Society*, 13(4), 212–214.
- Carrasco-Ortiz, H., Amengual, M., & Gries, S. (2019). Cross-language effects of phonological and orthographic similarity in cognate word recognition: The role of language dominance. *Linguistic Approaches to Bilingualism*, 1–29.
- Comesaña, M., Ferré, P., Romero, J., Guasch, M., Soares, A. P., & García-Chico, T. (2015). Facilitative effect of cognate words vanishes when reducing the orthographic overlap: The role of stimuli list composition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 41(3), 614–635.
- Comesaña, M., Soares, A., Sánchez-Casas, R., & Lima, C. (2012). Lexical and Semantic Representations in the Acquisition of L2 Cognate and Non-Cognate Words: Evidence from Two Learning Methods in Children. *British Journal of Psychology*, 103, 378–392.
- Dijkstra, T., Grainger, J., & Van Heuven, W. (1999). Recognition of Cognates and Interlingual Homographs: The Neglected Role of Phonology. *Journal of Memory and Language*, 41, 496–518.
- Dijkstra, T., Miwa, K., Brummelhuis, B., Sappeli, M., & Baayen, H. (2010). How Cross-Language Similarity and Task Demands Affect Cognate Recognition. *Journal of Memory and Language*, 62, 284–301.
- Dijkstra, T., & Van Heuven, W. J. B. (2002). The Architecture of the Bilingual Word Recognition System: From Identification to Decision. *Bilingualism: Language and Cognition*, 5(3), 175–197.
- Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2018). Multilink: A computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, 22(4), 657–679.
- Frances, C., Navarra-Barindelli, E., & Martin, C. D. (2021). Inhibitory and facilitatory effects of phonological and orthographic similarity on L2 word recognition across modalities in bilinguals. *Scientific Reports*, 11, 12812.
- Frances, C., Navarra-Barindelli, E., & Martin, C. D. (2022). Speaker Accent Modulates the Effects of Orthographic and Phonological Similarity on Auditory Processing by Learners of English. *Frontiers in Psychology*, 13, 1–12.
- Gerard, L. D., & Scarborough, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(2), 305–315.
- Gooskens, C., & Heeringa, W. (2004). Perceptive evaluations of Levenshtein dialect distance measurements using Norwegian dialect data. *Language Variation and Change*, 16(3), 189–207.
- Guasch, M., Boada, R., Ferré, P., & Sánchez-Casas, R. (2013). NIM: A Web-based Swiss Army knife to select stimuli for psycholinguistic studies. *Behavior Research Methods*, 45, 765–771.
- Izura, C., Cuetos, F., & Brysbaert, M. (2014). LexTALE-esp: A test to rapidly and efficiently assess the Spanish vocabulary size. *Psicológica*, 35(1), 49–66.
- Keuleers, E. (2013). vwr: Useful functions for visual word recognition research. R package version 0.3.0. <https://CRAN.R-project.org/package=vwr>
- Kondrak, G. (2000). A New Algorithm for the Alignment of Phonetic Sequences. *Proceedings of the 1st North American chapter of the Association for Computational Linguistics conference*, 288–295.
- Kondrak, G. (2001). Identifying Cognates by Phonetic and Semantic Similarity. *Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies*, 1–8.
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representation. *Journal of Memory and Language*, 33, 149–174.
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, 44, 325–343.
- Lenth, R. (2021). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.5.4. <https://CRAN.R-project.org/package=emmeans>
- Léwy, N., & Grosjean, F. (2008). The Léwy and Grosjean BIMOLA model. In F. Grosjean (Ed.), *Studying Bilinguals*, Oxford: Oxford University Press.
- Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for Comprehension of Speech. *Bilingualism: Language and Cognition*, 16(2), 304–324.

- Marian, V., & Spivey, M. (2003a). Bilingual and Monolingual Processing of Competing Lexical Items. *Applied Psycholinguistics*, 24, 173–193.
- Marian, V., & Spivey, M. (2003b). Competing Activation in Bilingual Language Processing: Within- and Between-Language Competition. *Bilingualism: Language and Cognition*, 6(2), 97–115.
- Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009). Language effects in second language learners and proficient bilinguals investigated with event-related potentials. *Journal of Neurolinguistics*, 22(3), 281–300.
- Midgley, K. J., Holcomb, P. J., & Grainger, J. (2011). Effects of Cognate Status on Word Comprehension in Second Language Learners: An ERP Investigation. *Journal of Cognitive Neuroscience*, 23(7), 1634–1647.
- Muntendam, A., van Rijswijk, R., Severijnen, G., & Dijkstra, T. (2022). The role of stress position in bilingual auditory word recognition: Cognate processing in Turkish and Dutch. *Bilingualism: Language and Cognition*, 1–12.
- Otwinowska, A., & Szewczyk, J. (2019). The more similar the better? Factors in learning cognates, false cognates and non-cognate words. *International Journal of Bilingual Education and Bilingualism*, 22(8), 974–991.
- Peeters, D., Dijkstra, T., & Grainger, J. (2013). The representation and processing of identical cognates by late bilinguals: RT and ERP effects. *Journal of Memory and Language*, 68(4), 315–332.
- Pivneva, I., Mercier, J., & Titone, D. (2014). Executive control modulates cross-language lexical activation during L2 reading: Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(3), 787–796.
- R Core Team (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>.
- Sanders, N. C., & Chin, S. B. (2009). Phonological Distance Measures. *Journal of quantitative linguistics*, 16(1), 96–114.
- Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in Europe as based on Levenshtein distance. *Bilingualism: Language and Cognition*, 15, 157–166.
- Schepens, J., Dijkstra, T., Grootjen, F., & Van Heuven, W. (2013). Cross-Language Distributions of High Frequency and Phonetically Similar Cognates. *PLoS ONE*, 8(5): e63006. doi:10.1371/journal.pone.0063006.
- Schwartz, A. I., Kroll, J. F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive Processes*, 22.
- Sherkina, M. (2003). The cognate facilitation effect in bilingual speech processing. *Toronto Working Papers in Linguistics*, 21, 135–151.
- Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for Comprehension of Speech. *Bilingualism: Language and Cognition*, 16(2), 304–324.
- Tercedor, M. (2010). Cognates as lexical choices in translation: interference in space-constrained environments. *Target*, 22(2), 177–193.
- Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, 9(4), 780–789.
- Van Heuven, W., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: a new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology* (2006), 67(6), 1176–1190.
- VanLangendonck, F., Peeters, D., Rueschemeyer, S. A., & Dijkstra, T. (2020). Mixing the stimulus list in bilingual lexicon decision turns cognate facilitation effects into mirrored inhibition effects. *Bilingualism: Language & Cognition*.
- Vitevitch, M. (2012). What do foreign neighbors say about the mental lexicon? *Bilingualism: Language and Cognition*, 15, 167–172. doi: 10.1017/S1366728911000149.
- Winther, I., Matussevych, Y., & Pickering, M. (2021). Cumulative frequency can explain cognate facilitation in language models. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 43.
- 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.