Cross-language effects of phonological and orthographic similarity in cognate word recognition

The role of language dominance

Haydee Carrasco-Ortiz, Mark Amengual and Stefan Th. Gries
Universidad Autonoma de Queretaro / University of California, Santa Cruz / University of California, Santa Barbara

This study investigated the extent to which phonological and orthographic overlap between the two languages of bilinguals predicts word processing abilities in their dominant and non-dominant languages. Forty-four English-dominant L1 English-L2 Spanish speakers and Spanish-dominant Spanish heritage speakers performed a lexical decision task while reading words in English and Spanish. We calculated orthographic and phonological similarity of cognate and noncognate words using the Levenshtein distance measure. Results showed that both bilingual groups benefited from orthographic similarity when reading Spanish and English words, whereas a facilitative effect was restricted to Spanish words that shared phonology across languages. These findings suggest a different contribution of phonological and orthographic similarity in bilingual word recognition, independently of language dominance.

Keywords: bilingual word recognition, phonological and orthographic similarities, language dominance

1. Introduction

Learning a second language (L2) appears to be easier when words in that language have similar forms and meaning with words in the first language (L1). This could be particularly advantageous for language learners whose both languages involve words with partial or complete written overlap and shared meaning (e.g. piano in Spanish and English). These translation equivalents with a high degree of orthographic, phonological, and semantic similarity across languages are defined

https://doi.org/10.1075/lab.18095.car | (proofs)
Linguistic Approaches to Bilingualism, **1–29**.
ISSN 1879–9264 | E-ISSN 1879–9272

as cognate words. This type of semantic and orthographic overlap is related to their common etymological origins. Cognate words can present different degrees of phonological overlap across languages. For instance, the English-Spanish cognate *angel* has a relatively different pronunciation in each language (/ˈeɪndʒəl/ vs /ˈaŋxel/), whereas a cognate word like *animal* has a much more similar pronunciation (/ˈænəməl/ vs /aniˈmal/). Although former bilingual studies on visual word recognition have consistently documented a processing advantage of cognate words (Dijkstra, Grainger, & van Heuven, 1999; Midgley, Holcomb, & Grainger, 2011; Peeters, Dijkstra, & Grainger, 2013), recent studies have shown that this facilitative effect of cognate words is modulated by the degree of orthographic and phonological overlap across both languages of the bilingual individual (Comesaña et al., 2015, 2012; Dijkstra, Miwa, Brummelhuis, Sappelli, & Baayen, 2010). The effects of variable degrees of orthographic and phonological similarity across languages have been mainly documented in studies testing cognate processing in the L2 where word recognition seems to be more sensitive to such cross-linguistic similarities.

A central question in bilingualism research has been to determine how bilinguals represent and use words from two distinct language systems (Costa & Sebastián-Gallés, 2014; Duñabeitia, Dimitropoulou, Dowens, Molinaro, & Martin, 2016; Kroll, Gullifer, & Rossi, 2013). Cognate words have been used to investigate the relative independence of the two languages in the bilingual mental lexicon and whether or not lexical access occurs only in the target language or in the context-relevant language (Dijkstra et al., 1999; Lemhöfer & Dijkstra, 2004; Lemhöfer, Dijkstra, & Michel, 2004). According to the language non-selective access hypothesis, bilinguals simultaneously activate word representations in both of their languages, even under conditions when only one language is required for a specific task (Dijkstra & van Heuven, 2002; Jared & Kroll, 2001; van Hell & de Groot, 1998; van Heuven, Dijkstra, & Grainger, 1998). A large number of behavioral studies have confirmed this hypothesis by showing faster and more accurate processing of cognate words in comparison with matched non-cognate control words. This cognate processing advantage has been reported across a variety of experimental tasks such as single word recognition tasks (Comesaña et al., 2015, 2012; Dijkstra et al., 1999; Lemhöfer & Dijkstra, 2004), translation tasks (de Groot, 1992; Sáchez-Casas, García-Albea, & Davis, 1992) and production tasks (Schwartz, Kroll, & Diaz, 2007). These studies have provided consistent empirical evidence suggesting that a cognate word can simultaneously activate its equivalent in the other language. However, these studies have provided less consistent results with respect to whether such parallel activation of overlapping orthographic and phonological units can occur while bilinguals are processing words in either of their L1 or L2. In the present study, we examine the extent to which processing of cognate and noncognate words is modulated by cross-language activation of orthographic and phonological units in both the dominant and non-dominant language.

2. Cross-language activation in the processing of cognates

Previous bilingual studies suggest that the cognate processing advantage is modulated by the degree of orthographic overlap across languages. Dijkstra et al. (2010) tested Dutch-English bilinguals performing an L2 English lexical decision task in which participants were presented with cognate words varying according to different degrees of orthographic overlap across English and Dutch. Based on a rating study, the selected cognates ranged on a continuum from identical cognates with a complete orthographic overlap across languages (e.g. *lamp-lamp*), to near-identical cognates with partial orthographic overlap (e.g. *flood*, translated as *vloed* in Dutch). Results showed a larger cognate processing advantage as orthographic similarity scores increased between Dutch and English cognate words. A similar processing advantage was also observed for identical cognate words that obtained higher scores of phonological similarity across languages. These results suggest that recognition of L2 cognates can simultaneously activate L1 cognate words at both the orthographic and phonological level of representation. Dijkstra et al. (2010) explained the cognate processing advantage as resulting from parallel co-activation of orthographic representations from both languages, leading to the activation of a common semantic representation.

While consistent empirical evidence of this cognate advantage in L2 contexts suggests that L2 cognate processing can benefit from its L1 equivalent, the effect of cognate status on L1 word processing appears to be more dependent on L2 proficiency (Midgley et al., 2011; Van Hell & Dijkstra, 2002). Midgley et al., (2011) conducted an event-related potential (ERP) study on English speaking learners of French while they read cognates and matched control words in English (L1) and French (L2) language blocks. Results showed different ERP responses for cognate and noncognate words in both language blocks. The N400 amplitude, which is assumed to reflect the ease with which the meaning of a word is processed, was more negative for control words in comparison to cognate words. The authors interpreted this reduced negativity on N400 amplitude as reflecting a processing advantage for cognate words that share meaning and orthography across a bilingual's two languages. Interestingly, Midgley et al. (2011) reported this processing facilitation when L1 English – L2 French bilinguals were reading in both of their languages, with more pronounced effects when bilinguals were reading cognates in their L2. That is, when compared to non-cognates, L2 cognate items showed later and longer differences in the N400 amplitudes than L1 cognate items. These processing differences were interpreted as reflecting more activation of L1 word forms while processing L2 cognates. Midgley et al. (2011) explained this cognate facilitation as an accumulation of the benefits of exposure to a given form – meaning association across two languages. Cognate facilitation could arise in the L1 and the L2 either via the partial activation of the orthographically similar translation equivalent in the case of close cognates or via shared whole-word orthographic representations in the case of identical cognates.

Cognate facilitation effects have also been reported while processing words in the dominant language. For instance, van Hell and Dijkstra (2002) investigated whether cognates from two non-dominant languages (L2 and L3) could influence word recognition in the dominant language (L1). Participants were two groups of trilinguals with different levels of L3 proficiency. In a lexical decision task, trilinguals read L1 (Dutch) words that were cognates with their translations in the L2 (English), cognates with their translations in the L3 (French) and L1 non-cognate words. Trilingual individuals who had more proficiency in the L3 showed a significant processing advantage for both L2 and L3 cognate words, whereas trilinguals with less proficiency in their L3 (French) showed faster recognition latencies to words that were cognates with English, but not to those that were cognates with French. These results indicate that L1 processing can be influenced by lexical knowledge from a less-dominant non-target language. However, a minimal level of exposure and proficiency in the bilinguals' non-dominant language is required to engender cognate effects in the dominant language.

Numerous hypotheses have been put forward to account for the observed advantage in processing cognate words compared to noncognate words. Based on a connectionist framework (Dijkstra et al., 2010; Midgley et al., 2011; Voga & Grainger, 2007) cognate words are assumed to have language-specific orthographic and phonological representations linked to a shared semantic representation. According to this account, cognate facilitation effects can be explained in terms of a high degree of orthographic, phonological and semantic overlap across L1 and L2 word translations. Thus, the presentation of a cognate word in one language leads to the activation of its cognate translation in the other language resulting in a more strongly activated semantic representation. Cognate facilitation reported in the L1 has been attributed to a greater exposure to the L2, which strengthens the associations between form-meaning across languages (Midgley et al. 2011).

The fact that cognate words can vary in terms of both overlapping orthographic and phonological units has been only sparsely addressed in earlier studies (Comesaña et al., 2015, 2012; Dijkstra et al., 2010; Schwartz et al., 2007). The gap in experimental work that examines the role of phonological similarity in cognate processing is remarkable given the considerable amount of empirical evidence showing how bilinguals can benefit from phonological overlapping units across languages during L2 and L3 word processing (Brysbaert, Van Dyck, & Van de Poel, 1999; Carrasco-Ortiz, Midgley, & Frenck-Mestre, 2012; Dimitropoulou, Duñabeitia, & Carreiras, 2011; Haigh & Jared, 2007; Lemhöfer & Dijkstra, 2004; Mulík, Carrasco-Ortiz, & Amengual, 2018). In an L2 production task, Schwartz, et al. (2007) found larger naming latencies when cognate words could be mapped onto two distinct pronunciations across languages (e.g. the different pronunciations of word 'base' in Spanish and English). Similarly, Dijkstra et al. (1999) reported delayed recognition latencies when cognate words did not share phonology while bilinguals were reading in their L2. Both studies suggest that the observed phonological effects arise because two distinct phonological representations associated to L1 and L2 cognate words are activated. However, it is still unclear how different degrees of consistency across phonological and orthographic units affect cognate word recognition. The question of whether cognate words with different degrees of phonological and orthographical overlap are equally represented in the bilingual lexicon and whether cross-language competition can arise when processing cognates both in the L1 and the L2 has not been entirely answered. The present study seeks to further investigate these questions with two groups of bilinguals that vary in terms of their age of exposure to their L2, their experience, and use of their two languages.

3. The present study

This study investigates whether cognate facilitation effects largely observed for cross-linguistic orthographic similarity also holds for cross-linguistic phonological similarity. More specifically, we examine how different degrees of orthographic and phonological overlap affect the recognition of cognate words as a function of language dominance. Participants consisted of two groups of English-Spanish bilinguals: native English speakers who were Spanish learners (English-dominant) and Spanish heritage speakers who had acquired English at an early age (Spanish-dominant). The selection of these two participant groups provides the opportunity to examine cross-linguistic influence in the processing of cognates with varying degrees of phonological and orthographic overlap by English-Spanish bilinguals, who are either English-dominant or Spanish-dominant, while maintaining the language pair constant. Participants performed a lexical decision task while reading words in two block lists: one block list included English translation equivalents of Spanish words, while the other block included the Spanish translations. This design used in a previous bilingual study (Midgley, et al. 2011), allowed us to compare the effect of orthographic and phonological similarity on cognate processing in both of the bilinguals' languages. Cognate and noncognate words varied in terms of orthographic and phonological similarity across Spanish and English. Using Levenshtein distance, we were able to obtain a continuous measure of phonological overlap ranging from a high degree of phonological overlap between English and Spanish cognate words (e.g. animal /ænəməl/ vs /animal/) to a low degree of phonological overlap across language (e.g. angel /eɪndʒəl/ vs /aŋxel/). Similarly, orthographic Levenshtein distance across languages was calculated for cognate and noncognate words. Control noncognate words had minimal or zero cross-language phonological and orthographic overlap, as most characters needed to be deleted in order to transform one English word (e.g. *cloud* /klaʊd/) into its equivalent in Spanish (e.g. *nube* /nuβe/).

On the basis of a previous study investigating bilingual cognate recognition (Dijkstra, et al. 2010), we expect cognate facilitation effects to be modulated by the degree of orthographic and phonological overlap across languages. As a result, reduced recognition latencies would be observed in cognate words with a high degree of both orthographic and phonological overlap because of the ease of processing overlapping lexical units. Previous studies have shown that cognate processing advantages can arise in all of the bilinguals' languages (Midgley et al., 2011; Van Hell & Dijkstra, 2002). Indeed, these studies have suggested that increasing language proficiency in L2 should lead to a greater exposure to the same association between an orthographic and semantic representation facilitating recognition of cognates relative to non-cognates. However, given the asymmetric connections between L1 and L2 word-forms hypothesized by connectionist models (Midgley, et al. 2011), we expect a cognate processing advantage to be modulated by language dominance in our bilingual groups. In other words, we expect a greater activation of cognate equivalents when bilinguals read words in their non-dominant language due to their stronger associative links with their translation equivalents in their dominant language than vice versa. This would result in a more attenuated cognate advantage effect in the bilinguals' dominant language. More specifically, Spanish-dominant bilinguals would benefit more from phonological and orthographic similarities when processing English cognates compared to Spanish cognates, while English-dominant bilinguals would show a greater cognate processing advantage while reading Spanish words compared to English words.

3.1 Method

3.1.1 *Participants*

Forty-eight participants were recruited to participate in the present study. The convenience sample consisted of early English-Spanish bilinguals (Spanish heritage speakers) and L1 English-L2 Spanish late bilinguals (L2 Spanish learners) who were undergraduate students enrolled in upper-division Spanish courses at the University of California, Santa Cruz. They received course credit for their participation. All participants reported normal speech and hearing and normal or corrected to normal vision.

All participants completed the Bilingual Language Profile (BLP) questionnaire (Birdsong, Gertken, & Amengual, 2012). The BLP is an instrument for assessing language dominance through self-reports and it produces a continuous dominance score and a general bilingual profile taking into account multiple dimensions: age of acquisition of the L1 and L2, frequency and contexts of use, competence in different skills, and attitudes towards each language. All of these factors are organized in four modules (language history, language use, language proficiency, and language attitudes), which received equal weighting (see Gertken, Amengual, & Birdsong, 2014). The BLP was administered prior to the beginning of the experiment, and was provided either in English or in Spanish, depending on participant preference. The classification of participants as Spanish-dominant or English-dominant was determined by their responses to the questionnaire, which generated a language particular score for each module, a global score for each language, and a global score of dominance. The point system was converted to a scale score with the Spanish score subtracted from the English score. Dominance scores ranged from −123.1 to 154.1. Negative values indicate Spanish-dominance whereas positive values indicate English-dominance. The data from four participants, whose dominance scores were closer to zero and therefore not clearly indicative of being dominant in one of their languages, were excluded from the analysis. As a result, the final sample included 22 early English-Spanish bilinguals (Spanish heritage speakers) and 22 L1 English-L2 Spanish late bilinguals (L2 Spanish learners).

The early English-Spanish bilinguals ($N = 22$) consisted of Spanish heritage speakers and individuals who had immigrated to the United States as children (19 female). These participants had been raised and educated in a bilingual environment in the United States and had extensive exposure to both Spanish and English on a daily basis. Their ages ranged from 18 to 41 ($M = 22$, $SD = 4.3$). The early bilinguals were from Generation 1.5 (G1.5) and Generation 2 (G2), following the categorization of Silva-Corvalán (1994). The G1.5 group ($N = 8$) consists of foreign-born (i.e., Mexico) individuals who arrived in the United States between the age of 6 and 11. The G2 group ($N = 14$) are individuals who were born in the United States with both of their parents born in Mexico. All participants in the Spanish heritage bilingual group were early sequential bilinguals who had been raised speaking Spanish exclusively at home, had acquired English during childhood, and were completing their education in the United States.

The English L1-Spanish L2 late bilinguals ($N = 22$) were L2 Spanish learners (19 female) who had been raised in a monolingual English household in the United

States, spoke English as their native language, and learned Spanish at school. Their ages ranged from 19 to 34 (M = 22.7, SD = 4.4). They reported not being native speakers or fluent in any other language. Each of these participants began their Spanish language study in high school and continued their education in Spanish at the university level. Participants from both groups were Spanish Studies majors, they had completed their education in the United States, and they were enrolled in upper-division Spanish courses at the time of testing. In order to be enrolled in upper-division Spanish courses, students are required to achieve a certain score on a placement test and advance beyond the six-quarter language course sequence to a set of core upper level courses in Spanish. Even though our participants did not complete a Spanish language proficiency test, all participants were enrolled in upper-division Spanish courses and had been required to achieve a certain score on a placement test and advance beyond the six-quarter Spanish language course sequence to a set of core upper-level courses in Spanish.

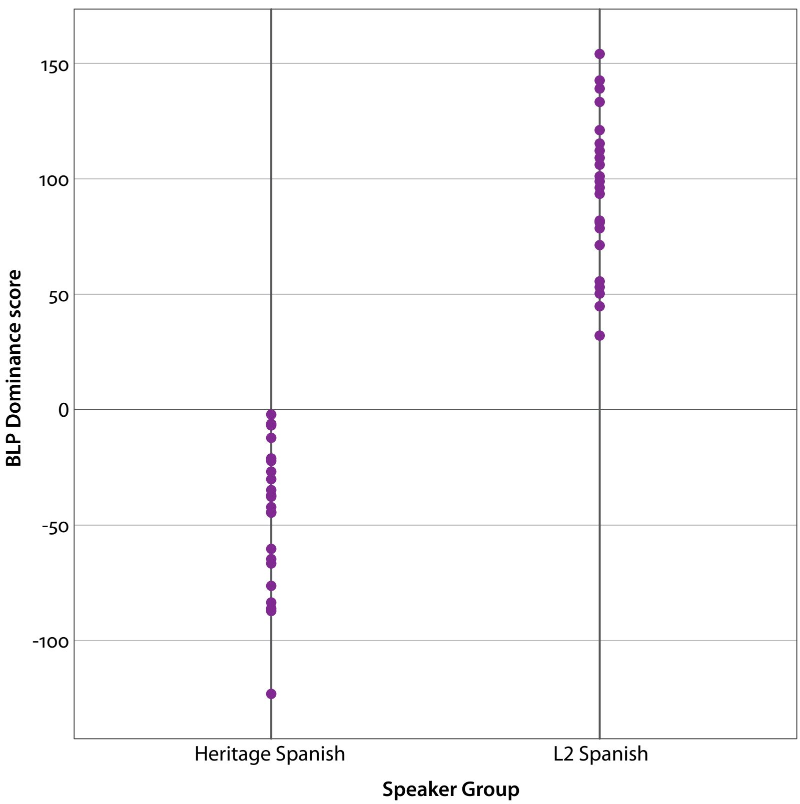

Figure 1. Language dominance as a function of speaker group according to the Bilingual Language Profile (BLP)

The mean dominance score for the Spanish heritage group was −46.1 ($SD = 31.1$) whereas the L2 Spanish learner group mean was 94.1 ($SD = 33.3$). Both groups differ from each other in terms of their language dominance scores ($t_{Welch} = 199.82$, $df = 8250.6$, $p < 0.0001$). Figure 1 provides the distribution of the Spanish heritage speakers and L2 Spanish learners as a function of their language dominance scores according to the BLP.

3.1.2 *Stimuli*

The stimulus materials consisted of 100 words that were cognates in Spanish and English and 100 control words that were noncognates between Spanish and English (see Appendix). Lexical properties between Spanish cognate and non-cognate words were controlled in terms of length in letters (mean = 5.38, $SD = 1$; mean = 5.53, $SD = 1.42$, respectively) and the number of orthographic neighbors calculated using the Levenshtein distance metric OLD20 (Yarkoni, Balota, & Yap, 2008) (mean = 1.54, $SD = .39$; mean = 1.55, $SD = .38$, respectively). Similarly, English cognate and noncognate words were controlled for length in letters (mean = 4.91, $SD = .84$; mean = 5.01, $SD = .75$, respectively) and the number of orthographic neighbors calculated using the Levenshtein distance metric OLD20 (mean = 1.80, $SD = .43$; mean = 1.71, $SD = .29$, respectively). Statistical analyses run for all these lexical properties revealed no significant differences between the mean values of cognate and noncognate words within each language (all p values > 0.19). In order to make log word frequency comparable across languages, we used the Zipf scale, which is a standardized measure of word frequency per million logged to the base of 10 + 3 (Van Heuven, Mandera, Keuleers & Brysbaert, 2014). Log word frequency in terms of Zipf values was controlled for Spanish cognate and noncognate words (mean = 4.67, $SD = .61$, mean = 4.49, $SD = .61$, respectively) and English cognate and noncognate words (mean = 4.70, $SD = .51$, mean = 4.64, $SD = .52$, respectively). These means were not statistically different (all $Fs < 1.90$, all p values < .16). Word frequencies per million and the number of orthographic neighbors were based on the EsPal database (Duchon, Perea, Sebastián-Gallés, Martí, & Carreiras, 2013) for Spanish words and the British Lexicon Project (Keuleers, Lacey, Rastle, & Brysbaert, 2011) and the English Lexicon Project databases (Balota et al., 2007) for English words.

Orthographic and phonological overlap between Spanish and English words was operationalized using the Levenshtein distance measure (Yarkoni et al., 2008). Orthographic and phonological overlap was computed based on the mean number of substitution, insertion, or deletion operations needed to turn a Spanish word into an English word (e.g., for *cost* and *costo*, Levenshtein distance value is 1 given the addition of "o"). Specifically, phonological Levenshtein distance calculation was based on the comparisons of phonological characters between the Spanish and English words (e.g., for *cost* /kast / and *costo* /kosto/ Levenshtein distance value is 2 given the substitution of "/a/" for "/o/" and the addition of /o/). Thus, phonological Levenshtein distance between the Spanish and English cognates was significantly smaller (mean = 3.34, *SD* = .92) to that between Spanish and English noncognates (mean = 5.40, *SD* = 1.12) ($t(198) = 14.19$, $p = .0001$). Similarly, orthographic Levenshtein distance between the Spanish and English cognates was significantly smaller (mean = .81, *SD* = .56) to that between Spanish and English noncognates (mean = 4.95, *SD* = 1.17) ($t(198) = 31.95$, $p = .0001$). Lexical properties and phonological and orthographic overlap are provided in Table 1.

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

Language	Word type	Phonological similarity	Orthographic similarity	Log frequency	Number of letters	Orthographic neighbors
Spanish	Cognate	3.34 (±.92)	.81 (±.56)	4.67 (±.61)	5.38 (±1)	1.54 (±.39)
	Noncognate	5.40 (±1.12)	4.95 (±1.17)	4.49 (±.61)	5.53 (±1.42)	1.55 (±.38)
English	Cognate	3.34 (±.92)	.81 (±.56)	4.70 (±.51)	4.91 (±.84)	1.80 (±.43)
	Noncognate	5.40 (±1.12)	4.95 (±1.17)	4.64 (±.52)	5.01 (±.75)	1.71 (±.29)

Experimental stimuli were split into two presentation lists such that each list contained 50 cognate and 50 noncognate words. Each list had two language versions, one in Spanish and one in English, so that all words had their translation equivalent in the other language. All word stimuli were counterbalanced across these four presentation lists to avoid repetition of the cognate words across languages. That is no one participant saw both the Spanish cognate and its English equivalent. Finally, a total of 20 Spanish non-words and 20 English non-words (16.7% of trials) served as probe items in the Spanish list version and the English list version, respectively.[1] The word stimuli in each presentation list were presented in a pseudorandom order.

3.1.3 *Apparatus and procedure*

Participants were seated comfortably in a sound-attenuated room. Word stimuli were displayed visually in white lowercase letters against a black background on a

1. Keeping the nonword ratio at 50/50 is desirable in order to minimize strategic processes that could potentially affect results (McNamara, 2005). However, a previous study showed that a nonword ratio below 50/50 seemed to be less likely to influence participants' use of strategic processes that could overestimate the reported lexical effects (Neely et al. 1989). Also, the amount of nonwords used in the present study is comparable with a previous ERP study (Geyer et al. 2011) on bilingual word recognition where the number of nonwords was around 20% of the trials.

computer screen that was positioned at approximately 1.5 m from the participant. The experiment was run on an Apple Macintosh computer. The procedure for each trial was as follows. First, a fixation cross appeared on the screen for 500 ms followed by 500 ms of black screen. Next, the target word was presented and remained on the screen until the participant responded. The onset of the participant's response was registered using a button box. Participants were instructed to read the words for meaning and performed a lexical decision task. They were asked to respond as quickly and as accurately as possible whether they were presented with a real word or a non-word. Reaction time (RT) was measured from the onset of the word stimulus. The order of the lists and languages was counterbalanced across participants such that all word stimuli were seen in both languages by a different participant. The experiment started with a short practice list (10 trials) to familiarize participants with the experimental procedure. Oral and written instructions were given in English and in Spanish according to the language of the stimulus list. There was a pause after completion of each list, the length of which was determined by the participant.

3.1.4 *Data analyses*

RTs of incorrect responses and those shorter than 100 msec and 2.5 SD above the participant mean (2.71% of all data) were removed from the analyses. Recognition latencies of the non-words, requiring a "no" response, were regarded as fillers and were not included in the analyses.

In order to investigate how both phonological and orthographic overlap affect visual word recognition in L1 and L2, we conducted a linear mixed-effects model analysis (using R's lme4::lmer) for the RTs obtained per item. The statistical analysis was informed by the following considerations regarding the variables' purpose:

- as the response, we considered RT (reaction time in ms);
- as predictors, we considered the variables PHON (phonological overlap across English and Spanish), ORTH (orthographic overlap across English and Spanish), LANGUAGE (English vs. Spanish), and GROUP (English-dominant vs. Spanish-dominant); the variable GROUP was rescaled into a numeric using the R function arm::rescale (see Gelman, 2008) such that −0.511 corresponds to English dominant and 0.489 to Spanish-dominant; (analyses with BLP did not change the results in any meaningful way, given the complete separation of values between the groups and the nearly perfectly bimodal distribution of values);
- as control variables, we used ZIPFFREQ (the word frequency per million logged to the base of 10 + 3) and LENGTH (their orthographic length in characters);

– as potential random effects, we considered SUBJECT (an id code for each experimental subject) and ITEM (the stimulus word).

Before we began with the statistical modeling, however, a thorough exploration of the data was undertaken to ensure their suitability for the subsequent analysis. This resulted in a variety of adjustments to the above design: First, and given its considerable skew, the response variable RT was inverted and multiplied by −1000 (as in Dijkstra et al. 2010). Secondly and unsurprisingly, we found a high correlation between PHON and ORTH: Pearson's $r = 0.782$ and each variable accounts for about 61% of the variability of the other. This is obviously something that needs to be addressed in order for the analysis to be able to achieve its main objective of determining whether PHON has an effect above and beyond ORTH. To that end, we applied a process called residualization to create two variables that might be called PHONwoutORTH and ORTHOLwoutPHON. As for the former, we fitted a generalized additive model (to capture even curvature effects and not just linear/straight-line effects) modeling PHON as a function of ORTH and then defined PHONwoutORTH as the residuals of that model; the reverse was done for ORTHOLwoutPHON.

4. Results

In a first step, we fit a linear mixed-effects regression in which we model the RT (transformed as mentioned above) as a function of GROUP, LANGUAGE, the residualized predictors just described, and all their pairwise interactions as fixed effects, LENGTH (as a polynomial to the degree of 2), ZIPFFREQ. and their interaction as fixed-effects control variables, and a maximal random-effects structure according to Barr, Levy, Scheepers, & Tily (2013) consisting of intercepts and GROUP slopes per ITEM and intercepts as well as slopes for PHONwoutORTH and ORTHOLwoutPHON and LANGUAGE per SUBJECT. With that model, we then proceeded, as in Dijkstra et al. (2010) again to eliminate outliers, which were defined as data points with standardized residuals <−2.5 or >2.5, which amounted to a loss of a mere 1.68% of the data points.

We then began a 3-step model selection process following Zuur et al. (2009): we (i) identified the best random-effects structure (using REML estimation), then (ii) narrowed down the fixed-effects structure (using ML estimation), and then (iii) computed the final model using REML again. The random-effects structure was trimmed down by likelihood ratio tests, the fixed-effects structure was trimmed down such that we removed effects from the regression model if both a likelihood ratio test and a comparison of AICc-values recommended deletion; at every model selection step we also checked for collinearity by computing variance inflation factors.

Trimming down the random-effects structure led to a model with no convergence problems and with only varying intercepts for ITEM and varying intercepts and slopes per SUBJECT, which we implemented for our exploration of the best fixed-effects structure of the model. The model with this random-effects structure was then re-fit using ML estimation and then explored with regard to (i) which predictors could be deleted and (ii) whether it was necessary to consider curved effects of PHONwoutORTH and ORTHOLwoutPHON. The final model arrived at following this process was then refitted with REML estimation and its highest-level predictors are shown in Table 2.

Table 2. Overview of the results of the final regression model

Predictors / controls	LR-stat.	df	Pdeletion
LANGUAGE : poly(PHONwoutORTH, 2)	4.934	2	0.085
GROUP : ORTHwoutPHON	5.681	1	0.017
GROUP : LANGUAGE	7.847	1	0.005
LANGUAGE : ORTHOLwoutPHON	6.37	1	0.012
poly(PHONwoutORTH, 2) : ORTHwoutPHON	12.476	2	0.002
LENGTH (control)	17.305	1	<0.001
ZIPFFREQ (control)	109.629	1	<0.001
Variance of intercepts per ITEM: 0.008	Variance of intercepts per SUBJECT: 0.031 Variance of LANGUAGE slopes per SUBJECT: 0.074		

The model's residuals gave no reason for concern and all variance inflation factors of the final model were ≤5.1. As is often the case in psycholinguistic experimentation and even more so with bilingual speakers, the model's explained variance is only moderate: R2 marginal = 0.164, R2 conditional = 0.444, and the random effects contribute more to the overall amount of variance explanation. However, the actual regression coefficients of such models are usually extremely hard to interpret, in particular if, as here, orthogonal polynomials and interactions are involved: in such cases, where curved effects might depend on other variables' levels, visualization is the only way to make sense of the results (much like GAM(M)s can only be interpreted visually). Thus, we are using effects plots of predicted reaction times to discuss the effects of the predictors our study is interested in; in all but one of the following plots, predicted transformed RTs are on the left y-axis and, for ease of comprehension, their corresponding raw RTs are on the right y-axis, the x-axis represents one predictor, and the other predictor is represented with colors as indicated by a legend.

For instance, Figure 2 shows the results for the interaction LANGUAGE: GROUP: The predictor LANGUAGE is on the x-axis and the predictor GROUP is indicated by colored lines as indicated by the legend.

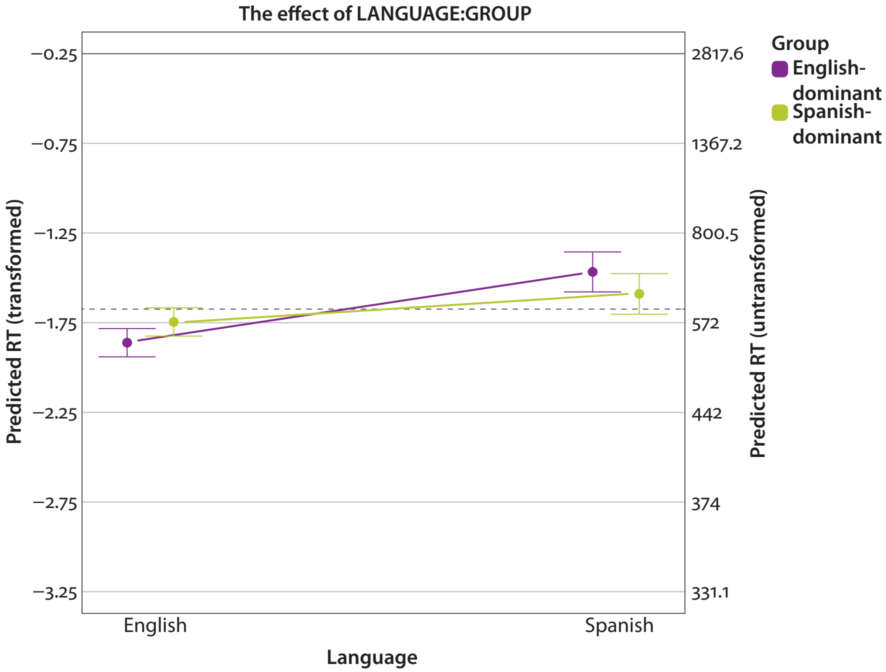

Figure 2. The interaction of GROUP and LANGUAGE

RTs are faster when LANGUAGE is English than when it is Spanish, but the size of the difference differs across groups: With the English – dominant speakers, the difference is more than twice as high (on the scale of the transformed RTs that was modeled) than with Spanish-dominant speakers; for better understanding, if we transform these RTs back to the raw values, we find that the difference between English and Spanish words for the English-dominant speakers is one of approximately 150 ms whereas that difference for the Spanish dominant speakers is only about 60 ms. Put differently, the English-dominant speakers are faster with English words and slower with Spanish words whereas the Spanish-dominant speakers are faster with Spanish words and slower with English words than the English-dominant speakers.

Figure 3 shows the other effect in which GROUP participates. The effect is essentially that increased orthographic dissimilarity slows English-dominant speakers down a bit more than Spanish-dominant speakers.

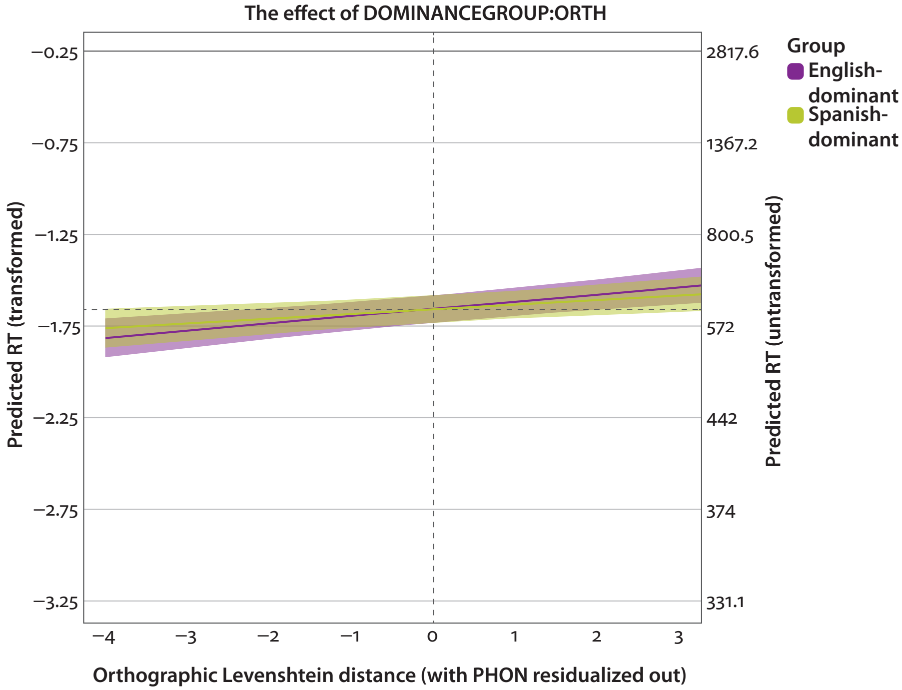

Figure 3. The interaction of GROUP and ORTHwoutPHON

Figure 4 shows the results for the interaction of LANGUAGE and poly(PHONwoutORTH, 2). As before, predicted RTs are on the y-axes, but now PHONwoutORTH is on the x-axis and LANGUAGE is indicated by the differently-colored lines. One can recognize the same kind of 'main effect' of LANGUAGE as in Figure 2: The response latencies for English words are smaller than those for Spanish words (the blue line is below the red line). But the (weak) interaction seems to be that English words are not sensitive to PHONwoutORTH (the blue line is virtually horizontal) whereas Spanish words are sensitive to PHONwoutORTH (the red line is rising). In other words, if phonological dissimilarity increases, speakers react increasingly more slowly to Spanish words, but equally to English words.

Figure 5 shows the results for the penultimate effect we are discussing, the interaction of LANGUAGE and ORTHwoutPHON. While the effect of ORTHwoutPHON does not involve curvature, its nature is quite similar to the one just discussed for Figure 4. When orthographic dissimilarity increases, speakers responded only a bit more slowly to English words, but quite noticeably more slowly to Spanish words.

Finally, there is a rather weak interaction PHONwoutORTH and ORTHwoutPHON, which is represented in a numeric version of a heatmap in

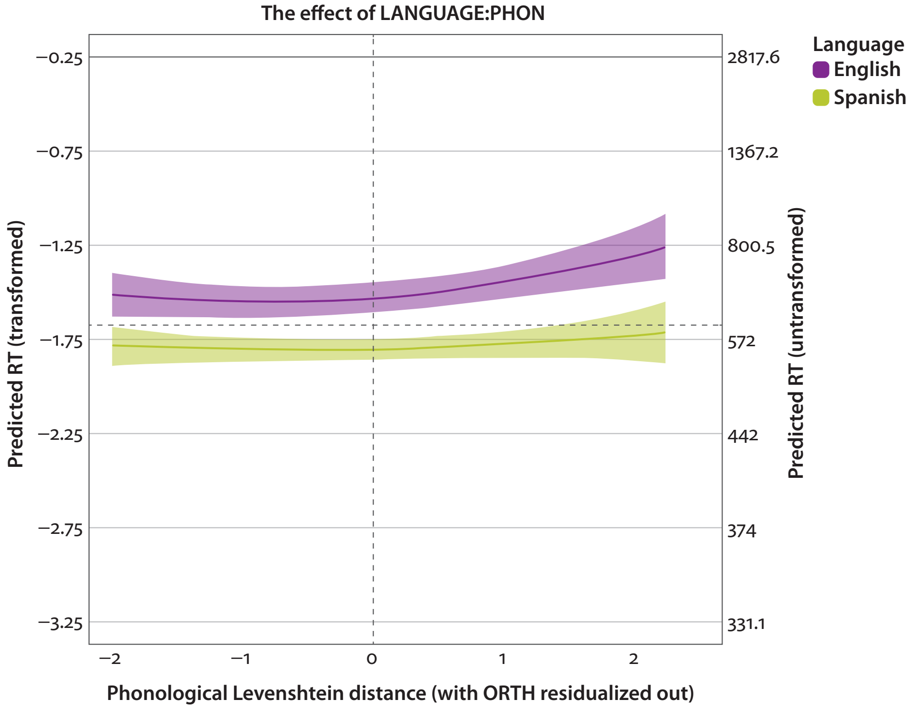

Figure 4. The interaction of LANGUAGE and PHONwoutORTH (2nd-degree polynomial)

Figure 6: PHONwoutORTH is on the x-axis, ORTHwoutPHON is on the y-axis, and the plotted numbers represent the predicted RT for each combination of values from the x- and y-axes. These numbers – 4 and 5 – were arrived at in the following way: We took the whole range of observed RTs and grouped them into their deciles labeled 0 to 9. The 4s and 5s plotted therefore indicate that for all combinations of PHONwoutORTH and ORTHwoutPHON the predicted RTs are in the middle two deciles of all RTs (hence the weakness of the effect). The nature of the effect is not easy to make sense of: subjects reacted faster when one of the distances was high(er) and the other was low(er), as is indicated by the 4s in the top left and the bottom right corner. In other words, participants were faster when phonological similarity was higher and orthographic similarity was lower and vice versa, i.e. when phonological similarity was lower and orthographic similarity was higher.

Finally, there are the two control variables, but to save space we are not visualizing those here. Their effects are entirely as predicted and completely linear: frequency speeds up RTs, length slows them down.

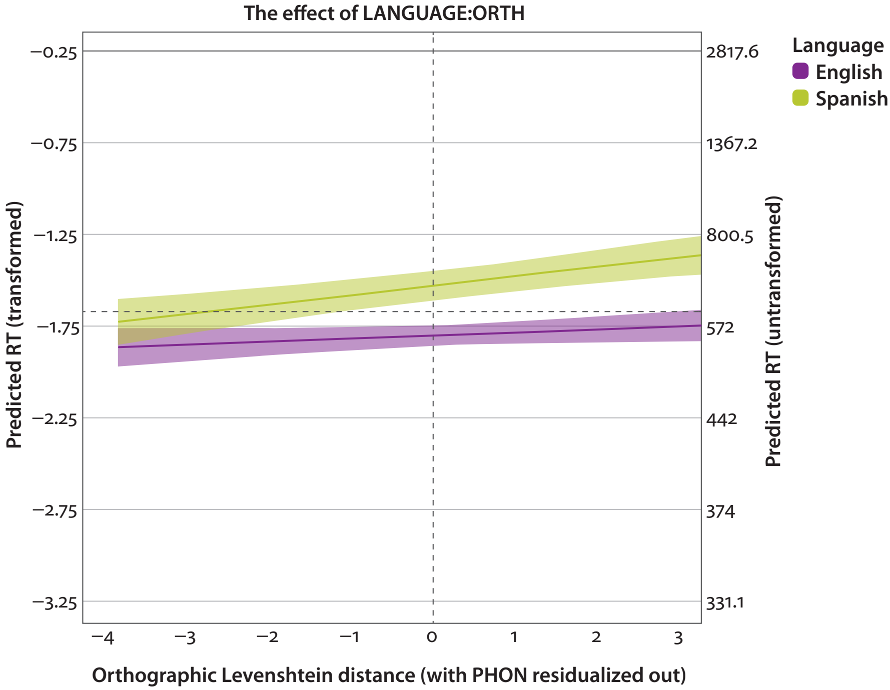

Figure 5. The interaction of LANGUAGE and ORTHwoutPHON

5. Discussion

The present study examined the extent to which cross-linguistic orthographic and phonological similarity predicts recognition of cognate and noncognate words as a function of bilinguals' language dominance. To this end, we presented cognate and noncognate words that varied in terms of orthographic and phonological similarity across English and Spanish languages based on the Levenshtein distance measure. Participants were native English speakers who had acquired Spanish later in life (English-dominant L2 Spanish learners) and early sequential Spanish-English bilinguals who had acquired Spanish and then English during childhood (Spanish-dominant Spanish heritage speakers). Both groups of speakers had been educated in the United States, were enrolled in upper-division Spanish courses, and were majoring in Spanish Studies at a public university. These participants performed a lexical decision task while reading two block lists of English and Spanish words.

As expected, English-dominant L1 English-L2 Spanish speakers were faster at recognizing English words than Spanish words, whereas Spanish-dominant Spanish heritage speakers were faster at recognizing Spanish words than English words (Figure 2). These results were informative with respect to the bilinguals' language dominance and confirmed what the BLP test had shown with the Spanish

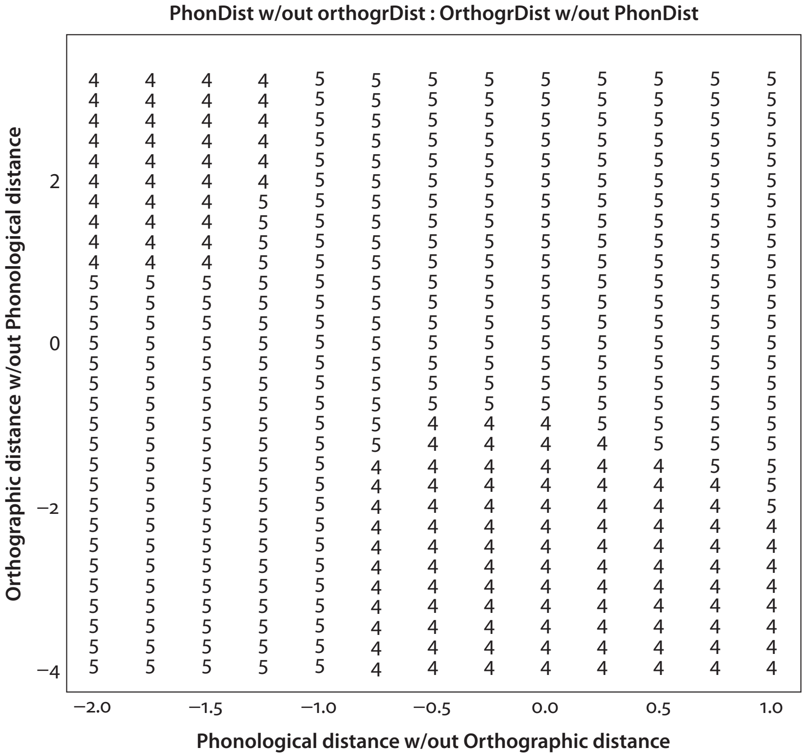

Figure 6. The interaction of PHONwoutORTH and ORTHwoutPHON

heritage speakers in this sample being more Spanish dominant and the L2 Spanish learners being more English dominant. More important for the present study, our results showed visual word recognition to be affected by cross-language orthographic and phonological similarity in both bilingual groups. Words with larger phonological similarity across languages decreased recognition latencies when bilinguals were reading in Spanish, whereas words with larger phonological overlap did not decrease recognition latencies when bilinguals were reading in English. Overall, English-dominant speakers appeared to be more sensitive to the degree of orthographic similarity in both languages compared to Spanish-dominant speakers. That is, both participant groups showed a facilitatory effect to words with larger orthographic similarity as they were reading in English and Spanish. These results indicate that bilinguals' lexical processing was modulated by the degree of both orthographic and phonological overlap across languages. Contrary to our hypothesis, the effect of orthographic and phonological similarity was not modulated by the bilinguals' language dominance.[2] Indeed, these cross-language effects of orthographic and phonological similarity in word recognition were true whether bilinguals were reading in their dominant or nondominant languages. As in a number of previous studies (Midgley et al., 2011; Mulík et al., 2018; Van Hell & Dijkstra, 2002), it appears that not only lexical properties in the dominant language can influence word recognition in the nondominant language, but also word recognition in the dominant language can be affected by lexical knowledge in the nondominant language. These results thus offer further evidence of nonselective language activation of both bilinguals' languages at the orthographic and phonological level of representation.

In line with previous studies (Dijkstra, 2010), a cognate facilitatory effect was observed to be modulated by the degree of orthographic overlap across languages. More importantly, our results showed this facilitatory recognition effect to be present in both languages of the bilingual individual. These results contrast with previous studies that failed to observe cognate effects in the bilinguals' dominant language (Caramazza & Brones, 1979; Gerard & Scarborough, 1989), but they are in line with a more recent electrophysiological study showing an influence of cognate status in visual word recognition in both of the bilinguals' languages (Midgley et al., 2011). Indeed, Midgley et al. (2011) reported a processing advantage (reduced N400) for cognate words when L1 English – L2 French bilinguals were reading in both their L1 and L2, with a more robust advantage when bilinguals were reading cognates in their L2. Interestingly, the facilitatory effect observed in the present study for words with larger orthographic overlap across languages appeared to be more prominent when participants were reading words in Spanish than in English. This can be due to the fact that even though English was not the dominant language for the Spanish heritage bilinguals, they were exposed to English cognates since early childhood. In fact, our results showed faster recognition latencies when Spanish heritage speakers were reading English words compared to Spanish words. This is also consistent with previous research showing heritage speakers' weaker performance in written tasks in their heritage language, specifically when compared to oral production tasks (Montrul, 2013). Furthermore, according to their responses to the BLP, both groups of speakers have had to read more in English than in Spanish throughout their educational experiences, and they also report a higher proficiency reading in English than

2. It is important to note that our participant groups, divided as a function of language dominance, also differed in terms of their age of acquisition. However, the results of this experiment also do not provide evidence that the age of acquisition and the differences in the group's learning environment modulate the effect of orthographic and phonological similarity in cognate word recognition.

in Spanish. A greater exposure to English cognates would have strengthened the associations between form-meaning across both languages (Midgley et al., 2011) and as a result these participants were more influenced by orthographic units in English while reading Spanish words in the lexical decision task.

In addition, visual recognition of cognate and noncognate words was sensitive to phonological similarity across languages. These findings provide evidence that phonological representations of both languages are activated when bilinguals are silently reading words in both of their dominant and nondominant languages. The reduced response latencies observed in both bilingual groups specifically for Spanish words with larger degree of phonological similarity is consistent with the hypothesis that facilitatory cognate effects increased with a high degree of phonological overlap between the pronunciations of two words (Comesaña et al., 2015, 2012; Dijkstra et al., 2010). Previous bilingual studies have also shown that facilitation can be obtained for interlingual word pairs that have phonological overlap, even in the absence of orthographic overlap (Carrasco-Ortiz et al., 2012; Dimitropoulou et al., 2011; Haigh & Jared, 2007; Thierry & Wu, 2004). Interestingly, the facilitatory effect observed for Spanish words with greater phonological similarity was not observed when bilinguals were reading their English translation equivalents. The inversed phonological effect observed for Spanish and English words is in line with other asymmetries reported in bilingual priming studies, where recognition of L2 cognate words was facilitated when they were preceded by L1 cognate translations, whereas recognition of L1 cognate words preceded by L2 cognate translations did not yield a facilitation (Midgley, Holcomb, & Grainger, 2009). These processing asymmetries in cognate recognition have been previously accounted for within different theoretical frameworks (Dijkstra et al., 2010; Kroll & Stewart, 1994; Kroll, van Hell, Tokowicz, & Green, 2010; Midgley et al., 2011; Voga & Grainger, 2007). According to connectionist accounts, the strength of associative links between L1 and L2 words are asymmetrical, which would predict more activation of L1 orthographic and phonological representations while processing L2 cognates than vice versa. It is thus likely that the presentation of English words had only partially activated the phonological representation of the Spanish translation equivalent, resulting in less interference from Spanish while enhancing processing of English words. In other words, it is possible that bilinguals' exposure and continued use of English on a daily basis may have made it less likely to activate the Spanish counterpart at the phonological level of representation while reading English words, mitigating cross-linguistic influence.

Overall, these results suggest that the effect of phonological similarity is dependent on orthographic overlap across languages. Facilitative effects were observed for words with larger phonological similarity but lower orthographic overlap and vice versa for words with lower phonological similarity but larger orthographic overlap. These results are comparable to previous studies reporting an independent contribution of phonological and orthographic overlap across languages. Indeed, facilitative effects were observed for words with substantial phonological overlap across languages but which have different spellings (Carrasco-Ortiz et al., 2012; Dimitropoulou et al., 2011; Haigh & Jared, 2007; Thierry & Wu, 2004). Similarly, previous findings suggest facilitative effects to words with higher degrees of orthographic overlap across languages, independently of the degree of phonological similarity (Dijkstra, et al. 2010). However, our results also showed that when orthographic similarity is held constant, the amount of phonological overlap affects word recognition and this could modify the direction of cognate effects (facilitative vs inhibitory). One model that can account for these results is the Bilingual Interactive Activation model (BIA+) proposed by Dijkstra and van Heuven (2002). According to this model, visual presentation of a cognate word leads to the co-activation of two different orthographic and phonological representations in each of the bilinguals' languages. Hence, a cognate processing advantage is expected to arise in a discontinuous manner depending on the degree of similarity across orthographic and phonological representations. Our results provide further evidence of this interplay between orthographic and phonological representations in cognate word recognition.

Furthermore, our results showed that participants responded more slowly to words with a high degree of orthographic and phonological similarity. It is possible that the frequency-of-use of both readings of cognates in each of the bilinguals' languages could have diminished the processing advantage of identical cognates, as suggested by previous studies (Peeters, et al., 2013; Gollan, Montoya, Cera & Sandoval, 2008). Peeters et al. (2013) found that lexical frequency of L1 cognates can modulate the processing advantage of L2 identical cognate. Gollan et al. (2008) suggest that bilinguals' use of words is distributed over two languages, so the more they use an identical cognate in one language, the less they use the same word in the other language. Thus, recognition latencies should be slower when identical cognates are read in the less frequently used language in comparison to the more frequently used language. However, further work considering frequency of use of identical cognates is necessary to elucidate this effect.

While our data clearly show sensitivity to the degree of orthographic and phonological similarity across languages when participants were reading in both of their languages, we do not negate the possible influence of the stimulus list composition as the number of nonwords was less than that reported in previous studies (Comesaña, et al. 2015; Dijkstra, et al. 2010). Previous research on semantic priming suggests that different sizes of the nonword ratio can be associated with the use of more automatic or strategic cognitive processes (McNamara, 2005; Neely et al. 1989). Specifically, larger proportions of nonwords over words in lexical decision tasks may lead to strategic process that could potentially increase the observed lexical effects. However, these strategic processes are less likely to occur when the number of nonwords is smaller with respect to words as it was the case in our study. Also, previously reported cognate effects, independently of their direction (facilitatory vs inhibitory), have been found to be persistent across a variety of experimental designs and tasks that a small proportion of nonwords in our study could not be considered the main driving factor behind our results. Hence, while we cannot make the claim that a small proportion of nonwords in our stimulus lists have had played a role in our results, we think it is safe to say that the degree of orthographic and phonological overlap across languages was indeed the major factor.

6. Conclusion

The present study provides evidence that word processing in both the dominant and non-dominant language is modulated by orthographic and phonological similarity across languages. A facilitative effect was observed to be larger as orthographic overlap across languages increased, especially when participants were reading Spanish words. Similarly, the effect of phonological similarity seems to be more dependent on the bilinguals' English language environment as English phonological representations were more strongly activated during the recognition of Spanish words than in the opposite direction. These findings are in line with connectionist models suggesting a different contribution of shared orthographic and phonological representations in bilingual word recognition.

References

Baayen, R. H. (2010). Demythologizing the word frequency effect: A discriminative learning perspective. *The Mental Lexicon*, 5(3), 436–461.

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., … Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, 39(3), 445–459. https://doi.org/10.3758/BF03193014

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of memory and language*, 68(3). doi: https://doi.org/10.1016/j.jml.2012.11.001

Brysbaert, M., Van Dyck, G., & Van de Poel, M. (1999). Visual word recognition in bilinguals: evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance*, 25(1), 137–148.

Caramazza, A., & Brones, I. (1979). Lexical access in bilinguals. *Bulletin of the Psychonomic Society*, 13(4), 212–214. https://doi.org/10.3758/BF03335062

Carrasco-Ortiz, H., Midgley, K. J., & Frenck-Mestre, C. (2012). Are phonological representations in bilinguals language specific? An ERP study on interlingual homophones. *Psychophysiology*, 49(4), 531–543. https://doi.org/10.1111/j.1469-8986.2011.01333.x

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. https://doi.org/10.1037/xlm0000065

Comesaña, M., Sánchez-Casas, R., Soares, A. P., Pinheiro, A. P., Rauber, A., Frade, S., & Fraga, I. (2012). The interplay of phonology and orthography in visual cognate word recognition: An ERP study. *Neuroscience Letters*, 529(1), 75–79. https://doi.org/10.1016/j.neulet.2012.09.010

Costa, A., & Sebastián-Gallés, N. (2014). How does the bilingual experience sculpt the brain? *Nature Reviews Neuroscience*, 15(5), 336–345. https://doi.org/10.1038/nrn3709

de Groot, A. M. B. (1992). Bilingual Lexical Representation: A Closer Look at Conceptual Representations. In En R. Frost & L. Katz (Eds.), *Advances in Psychology* (Vol. 94, pp. 389–412). North-Holland. https://doi.org/10.1016/S0166-4115(08)62805-8

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(4), 496–518. https://doi.org/10.1006/jmla.1999.2654

Dijkstra, T., Miwa, K., Brummelhuis, B., Sappelli, M., & Baayen, H. (2010). How cross-language similarity and task demands affect cognate recognition. *Journal of Memory and Language*, 62(3), 284–301. https://doi.org/10.1016/j.jml.2009.12.003

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. https://doi.org/10.1017/S1366728902003012

Dimitropoulou, M., Duñabeitia, J. A., & Carreiras, M. (2011). Masked translation priming effects with low proficient bilinguals. *Memory cognition*, 39(2), 260–275. https://doi.org/10.3758/s13421-010-0004-9

Duchon, A., Perea, M., Sebastián-Gallés, N., Martí, A., & Carreiras, M. (2013). EsPal: One-stop shopping for Spanish word properties. *Behavior Research Methods*, 45(4), 1246–1258. https://doi.org/10.3758/s13428-013-0326-1

Duñabeitia, J. A., Dimitropoulou, M., Dowens, M. G., Molinaro, N., & Martin, C. (2016). The Electrophysiology of the Bilingual Brain. In En R. R. Heredia, J. Altarriba, & A. B. Cieślicka (Eds.), *Methods in Bilingual Reading Comprehension Research* (pp. 265–312). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4939-2993-1_11

Gelman, Andrew. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, 27, 2865–2873. https://doi.org/10.1002/sim.3107

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. https://doi.org/10.1037/0278-7393.15.2.305

Geyer, A., Holcomb, P. J., Midgley, K. J., & Grainger, J. (2011). Processing words in two languages: An event-related brain potential study of proficient bilinguals. *Journal of Neurolinguistics*, 24(3), 338–351. https://doi.org/10.1016/j.jneuroling.2010.10.005

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, 58, 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Haigh, C. A., & Jared, D. (2007). The activation of phonological representations by bilinguals while reading silently: evidence from interlingual homophones. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 33(4), 623–644. https://doi.org/10.1037/0278-7393.33.4.623

Jared, D., & Kroll, J. F. (2001). Do Bilinguals Activate Phonological Representations in One or Both of Their Languages When Naming Words? *Journal of Memory and Language*, 44(1), 2–31. https://doi.org/10.1006/jmla.2000.2747

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2011). The British Lexicon Project: Lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavior Research Methods*, 44(1), 287–304. https://doi.org/10.3758/s13428-011-0118-4

Kroll, J. F., Gullifer, J. W., & Rossi, E. (2013). The Multilingual Lexicon: The Cognitive and Neural Basis of Lexical Comprehension and Production in Two or More Languages. *Annual Review of Applied Linguistics*, 33, 102–127. https://doi.org/10.1017/S0267190513000111

Kroll, J. F., & Stewart, E. (1994). Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections Between Bilingual Memory Representations. *Journal of Memory and Language*, 33(2), 149–174. https://doi.org/10.1006/jmla.1994.1008

Kroll, J. F., van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The Revised Hierarchical Model: A critical review and assessment. *Bilingualism*, 13(3), 373–381. https://doi.org/10.1017/S136672891000009X

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: effects of code similarity in language-specific and generalized lexical decision. *Memory cognition*, 32(4), 533–550. https://doi.org/10.3758/BF03195845

Lemhöfer, K., Dijkstra, T., & Michel, M. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes*, 19(5), 585–611. https://doi.org/10.1080/01690960444000007

McNamara, T. P. (2005). *Essays in cognitive psychology. Semantic priming: Perspectives from memory and word recognition*. New York, NY, US: Psychology Press. https://doi.org/10.4324/9780203338001

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. https://doi.org/10.1016/j.jneuroling.2008.08.001

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. https://doi.org/10.1162/jocn.2010.21463</p>

Montrul, S. (2013). How "native" are heritage speakers?. *Heritage Language Journal*, 10(2), 153–177.

Mulík, S., Carrasco-Ortiz, H., & Amengual, M. (2018). Phonological activation of first language (Spanish) and second language (English) when learning third language (Slovak) novel words. *International Journal of Bilingualism*, doi: https://doi.org/10.1177/1367006918781061

Neely, J. H., Keefe, D. E., & Ross, K. L. (1989). Semantic priming in the lexical decision task: roles of prospective prime-generated expectancies and retrospective semantic matching. Journal of Experimental Psychology. *Learning, Memory, and Cognition*, 15(6), 1003–1019. https://doi.org/10.1037/0278-7393.15.6.1003

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. https://doi.org/10.1016/j.jml.2012.12.003

Sáchez-Casas, R. M., García-Albea, J. E., & Davis, C. W. (1992). Bilingual lexical processing: Exploring the cognate/non-cognate distinction. *European Journal of Cognitive Psychology*, 4(4), 293–310. https://doi.org/10.1080/09541449208406189

Schwartz, A. I., Kroll, J. F., & Diaz, Mi. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive Processes*, 22. https://doi.org/10.1080/01690960500463920

Thierry, G., & Wu, Y. J. (2004). Electrophysiological evidence for language interference in late bilinguals: *NeuroReport*, 15(10), 1555–1558. https://doi.org/10.1097/01.wnr.0000134214.57469.c2

van Hell, J. G., & de Groot, A. M. B. (1998). Conceptual Representation in Bilingual Memory: Effects of Concreteness and Cognate Status in Word Association. *Bilingualism: Language and Cognition*, 1(3), 193–211. https://doi.org/10.1017/S1366728998000352

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. https://doi.org/10.3758/BF03196335

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic Neighborhood Effects in Bilingual Word Recognition. *Journal of Memory and Language*, 39(3), 458–483. https://doi.org/10.1006/jmla.1998.2584

van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: a new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology* (2006), 67(6), 1176–1190. https://doi.org/10.1080/17470218.2013.850521

Voga, M., & Grainger, J. (2007). Cognate status and cross-script translation priming. *Memory & Cognition*, 35(5), 938–952. https://doi.org/10.3758/BF03193467

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

Zuur, A., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). *Mixed Effects Models and Extensions in Ecology with R*. New York: Springer-Verlag. https://doi.org/10.1007/978-0-387-87458-6

Appendix

Cognate words		Noncognate words	
Spanish	**English**	**Spanish**	**English**
ácido	acid	consejo	advice
activo	active	financiar	fund
actor	actor	manzana	apple
actual	actual	lote	batch
aire	air	batalla	battle

(*continued*)

Cognate words		Noncognate words	
Spanish	**English**	**Spanish**	**English**
alarma	alarm	playa	beach
ángel	angel	rayo	beam
animal	animal	frijol	bean
área	area	cerveza	beer
árido	arid	baya	berry
aspecto	aspect	llama	blaze
banco	bank	cloro	bleach
base	base	cerebro	brain
básico	basic	marca	brand
benigno	benign	pan	bread
cable	cable	puente	bridge
calma	calm	brillante	bright
cámara	camera	café	brown
canoa	canoe	entierro	burial
caso	case	ráfaga	burst
casual	casual	arbusto	bush
causa	cause	llamada	call
clase	class	silla	chair
clínica	clinic	cambio	change
color	color	reloj	clock
costo	cost	nube	cloud
crédito	credit	frio	cold
cubo	cube	peine	comb
cura	cure	algodón	cotton
curva	curve	tos	cough
debate	debate	peligro	danger
dieta	diet	muerte	death
directo	direct	cena	dinner
disco	disc	platillo	dish
dolar	dollar	cajón	drawer
drama	drama	sueño	dream
élite	elite	vestido	dress
error	error	bebida	drink

Cognate words		Noncognate words	
Spanish	**English**	**Spanish**	**English**
escape	escape	temprano	early
evento	event	tierra	earth
familia	family	vacío	empty
figura	figure	otoño	fall
formal	formal	granja	farm
guitarra	guitar	sentir	feel
idea	idea	campo	field
ideal	ideal	pelea	fight
imagen	image	pez	fish
insecto	insect	carne	flesh
límite	limit	vuelo	flight
línea	line	piso	floor
león	lion	espuma	foam
liquido	liquid	bosque	forest
lista	list	congelar	freeze
local	local	juego	game
lógica	logic	ajo	garlic
mapa	map	fantasma	ghost
marco	mark	gris	gray
mérito	merit	verde	green
metal	metal	mano	hand
modo	mode	salud	health
modelo	model	cielo	heaven
momento	moment	altura	height
motor	motor	miel	honey
músculo	muscle	caballo	horse
música	music	hierro	iron
nación	nation	riñon	kidney
océano	ocean	amable	kind
órgano	organ	beso	kiss
palma	palm	golpe	knock
pánico	panic	aprender	learn
papel	paper	nivel	level

(*continued*)

Cognate words		Noncognate words	
Spanish	**English**	**Spanish**	**English**
parte	part	préstamo	loan
pasta	pasta	piedad	mercy
foto	photo	feliz	merry
piano	piano	leche	milk
pino	pine	espejo	mirror
planta	plant	mes	month
plato	plate	boca	mouth
poeta	poet	angosto	narrow
público	public	nueve	nine
puré	pure	ruido	noise
radio	radio	orgullo	pride
raro	rare	leer	read
real	real	grito	scream
renta	rent	forma	shape
rosa	rose	camisa	shirt
rural	rural	falda	skirt
secreto	secret	humo	smoke
signo	sign	canción	song
sólido	solid	piedra	stone
suma	sum	fuerte	strong
símbolo	symbol	verano	summer
terror	terror	sudor	sweat
texto	text	hilo	thread
tono	tone	garganta	throat
tren	train	pueblo	town
tubo	tube	feo	ugly
vacante	vacant	ventana	window
víctima	victim	año	year
vital	vital	joven	young

Address for correspondence

Haydee Carrasco-Ortiz
Associate professor
Facultad de Lenguas y Letras
Universidad Autonoma de Queretaro
Circuito Fray Junípero Serra Km 8
Queretaro, 76140
Mexico

haydee.carrasco@uaq.edu.mx
https://orcid.org/0000-0002-2776-5965

Co-author details

Mark Amengual
University of California, Santa Cruz

amengual@ucsc.edu
https://orcid.org/0000-0003-2007-9687

Stefan Th. Gries
University of California, Santa Barbara

stgries@linguistics.ucsb.edu
https://orcid.org/0000-0002-6497-3958

Publication history

Date received: 2 November 2018
Date accepted: 17 July 2019

Author queries

Please provide a complete reference for the citation '(Dijkstra, 2010)' in this article.

Please provide a citation for the reference id "CIT0001 (Baayen, R. H. (2010))" since citation is missing in the article.