Original Article

The International Journal of Bilingualism

International Journal of Bilingualism
1–21

DOI: 10.1177/13670069211033359
journals.sagepub.com/home/ijb
SAGE

Cognate facilitation effect during auditory comprehension of a second language: A visual world eye-tracking study

Filip Andras
Marta Rivera
Teresa Bajo
University of Granada, Spain

Paola E Dussias
The Pennsylvania State University, USA

Daniela Paolieri
University of Granada, Spain

Abstract

Aims and Objectives: The cognate facilitation effect (CFE) is a robust effect in language production and visual word comprehension, but evidence for CFE during auditory comprehension is still scarce. This study aimed to explore the CFE during auditory comprehension of a second language (L2) while manipulating proficiency in the L2 and cognate type. These two variables are known to influence the CFE.

Methodology: Low and highly proficient Spanish–English bilinguals listened to individual words in their L2, English, that shared high, low, and no phonological overlap (PO) with their native language Spanish. We designed a visual world paradigm task that consisted of selecting an image shown as a spoken word unfolded in time while eye movements were recorded.

Data and Analysis: Response times revealed a clear CFE in low proficiency bilinguals, while this effect was absent in highly proficient bilinguals. The eye-tracking (ET) data showed late coactivation of low-PO words and, surprisingly, no coactivation of high-PO words in low proficiency bilinguals. Highly proficient bilinguals showed no clear pattern of language coactivation in the ET data. The English monolingual control group showed no effects during the critical time window.

Conclusions: These results are interpreted within the framework of L2 processing models. At low levels of proficiency, the PO between translations facilitates access to meaning. On the other

Corresponding author:
Daniela Paolieri, Mind, Brain, and Behavior Research Centre (CIMCYC), University of Granada, Campus de Cartuja s/n, Granada 180 71, Spain.
Email: dpaolieri@ugr.es

hand, highly proficient bilinguals no longer benefit from the PO between translations, at least for concrete and simple nouns.
Originality: The findings demonstrate a clear CFE in auditory comprehension. Proficiency in L2 and PO modulated the effect, as shown in both the response time and in the ET data, respectively.
Implications: These findings suggest that at low levels of L2 proficiency, learners more easily access the conceptual information if the auditory input is similar to their native language. Nevertheless, as proficiency increases, this facilitation disappears.

Keywords
Cognate facilitation effect, visual world paradigm, language non-selectivity, lexical access in bilinguals, second language auditory comprehension

Introduction

An important finding in bilingual research is that bilinguals maintain both of their languages active even in completely unilingual settings (for reviews see Kroll & Bialystok, 2013; Kroll et al., 2014, 2018). Language coactivation has been studied with unique words such as cognates. Cognates are words that share the same meaning and overlap phonologically/orthographically between languages. Words such as *piano* and *radio* are perfect examples of English–Spanish orthographic cognates because although their pronunciation differs across both languages, their spelling remains the same. Similarly, terms such as dentist–*dentista* or balcony–*balcón* are considered English–Spanish cognates as well, even though their orthographic overlap is not perfect. Crucially, bilinguals process cognates faster than noncognates, and this effect has been termed the *cognate facilitation effect* (CFE). This effect serves as a proxy of language coactivation. The CFE has been studied in language production and language comprehension using a variety of language pairs, and researchers have investigated the phenomenon in both first language (L1) and second language (L2) contexts. Although the CFE has been studied through language comprehension, most evidence comes from production studies (Costa et al., 2000; see also Hoshino & Kroll, 2008; Jared et al., 2012). For example, during a lexical decision and progressive demasking task, Dutch–English bilinguals exhibited shorter response times for cognates rather than noncognates (Dijkstra et al., 1999).

Similarly, bilingual adults and professional translators exhibited shorter reading times when reading for translation but not for repetition (Macizo & Bajo, 2006). The available findings taken together, indicate that the CFE is a robust effect found in language production, visual word recognition, and reading. However, the evidence for the CFE during auditory comprehension is still scarce (see, Blumenfeld & Marian, 2007; Valente et al., 2018). The purpose of the present study is to provide further evidence for CFE in L2 auditory comprehension and clarify the conditions under which CFE manifests. We focus on two variables that are known to influence the magnitude of the CFE: proficiency in L2; and the degree of phonological overlap (PO) between translation equivalents. With this purpose, we will use the visual world paradigm alongside eye-tracking (ET) methodologies to study language coactivation during auditory comprehension in the bilingual's L2 (Tanenhaus et al., 1995, 2000).

Current bilinguals processing models such as BIA-d (Grainger et al., 2010) or RHM (Kroll & Stewart, 1994; Kroll et al., 2010) take developmental perspectives and include specific predictions about the language dynamics that take place as proficiency in L2 is acquired. However, coactivation effects during L2 auditory comprehension have been mainly studied with highly proficient bilinguals (Blumenfeld & Marian, 2007; Ju & Luce, 2004; Marian & Spivey, 2003a, 2003b; Spivey

& Marian, 1999; Weber & Cutler, 2004) and to a lesser extent with a population at early stages of L2 acquisition (e.g., Valente et al., 2018). During the initial language learning stages, associations at the lexical level need to be made between L1 and L2 to access the semantic–conceptual level. As such, these models predict that low proficient bilinguals would be slower overall during L2 processing because their L1 mediates access to the L2 word form. Following increased proficiency in their L2, comprehension for fluent bilinguals would be relatively effortless due to the establishment of direct lexical–conceptual links, this transition would be similar for both production and comprehension as stated explicitly in the BIA-d model.

On the other hand, these models (BIA-d, Grainger et al., 2010; RHM, Kroll & Stewart, 1994; Kroll et al., 2010; and BIA+, Dijkstra & van Heuven, 2002; Dijkstra et al., 2010) either explicitly or implicitly make predictions about the degree of form overlap as well. At initial stages of L2 acquisition, the RHM postulates that the lexical level associative links should be stronger for cognates because of the greater orthographic overlap/PO. Consequently, this similarity across the languages facilitates access to the semantic representation, and the BIA+ posits that the amount of facilitation should be proportional to the amount of formal overlap.

Empirical evidence regarding PO concerning cognate facilitation and language coactivation has been mostly explored in visual word recognition studies (e.g., Comesaña et al., 2012, 2015; Dijkstra et al., 1999, 2010), and to a lesser extent in word naming (Schwartz et al., 2007; see also Iniesta et al., under review; for word writing evidence) and spoken language comprehension (Valente et al., 2018, see also, Blumenfeld & Marian, 2007). In visual word recognition, findings are somewhat mixed, and depend on the task and whether the target word is an orthographically identical cognate or a non-identical cognate. For example, Dijkstra et al. (2010) found a facilitation effect for PO during a lexical decision task only for identical cognates and no effect during a language decision task. Furthermore, there seems to be an interplay between orthography and phonology regarding cognates in visual word comprehension (e.g., Comesaña et al., 2012, 2015; Schwartz et al., 2007).

In spoken language comprehension, Blumenfeld and Marian (2007) used a visual world paradigm to test both German-native speakers immersed in an English context and English-native speakers who were learning German. The authors compared the proportion of fixations to distractors that were phonologically related and unrelated to either noncognate or cognate target word. For example, in noncognate trials, for the target ***desk***–*Schreibstisch*, a higher proportion of fixations was expected from the phonologically related interlingual distractor ***lid***–*Deckel* in comparison to the phonologically unrelated distractor. Importantly, the authors manipulated the PO between targets and distractors reaching low, medium, and high overlap between the initial portion of the target words. Results showed that, when listening to noncognates, only German native speakers looked more at the interlingual distractors. On the other hand, when presented with cognates, both German and English native speakers looked more at the interlingual distractors. Therefore, presenting cognates as targets seemed to boost language co-activation. Also, the degree of overlap increased the coactivation of German distractors, but again, only when the target words were cognates. Moreover, this coactivation was shorter for German native speakers, suggesting that this group had resolved the language competition earlier unlike English native speakers that were learning German as their L2.

Moreover, in a recent study, Valente et al. (2018) explored the acquisition of cognate words in children and adults and subsequently tested them on auditory recognition task and a go/no-go auditory lexical decision task. Importantly, the authors manipulated the degree of phonological and orthographic overlap for cognates yielding three conditions: cognates with high phonological and high orthographic overlap; cognates with high phonological and low orthographic overlap; and cognates with low phonological and low orthographic overlap. In the auditory recognition task, the authors found a CFE for both adults and children. In the lexical decision task, the authors found cognate facilitation in one measure (latencies) for adults and in two measures (latencies and error rates) for children. Their data also showed that only children were sensitive to the phonological and orthographical overlap of cognates while adults showed no differences between the three conditions. Concretely, in the lexical decision task, children showed a linear pattern with the fastest latencies for the most overlapping condition, slower latencies for the condition with the mismatch between phonological (high) and orthographical (low) overlap, and the slowest latencies for the least overlapping condition. In the error rate analyses of the lexical decision task and the auditory recognition task, the pattern was similar, although not perfectly linear. All in all, there was overall cognate facilitation for both children and adults but only children were sensitive to the different amount of orthographical overlap and PO in cognates.

In sum, compared to visual recognition, the two auditory comprehension studies suggest that PO is positively related to language coactivation during auditory comprehension (although it was restricted to children in the study by Valente et al., 2018).

The present study

The present study aimed to explore how proficiency in L2 (English) and cognate type or the degree of PO modulated the CFE during auditory comprehension in a visual world paradigm task (Tanenhaus et al., 1995, 2000). In this task, Spanish–English bilinguals with low and high proficiency in L2 (English) listened to single English words that shared high (high-PO), low (low-PO), and no PO (no-PO) with its Spanish translation. The task consisted of selecting an image, as fast and as accurately as possible, that corresponded to the target spoken word while eye movements were recorded. An advantage of using the visual world procedure is that it helps in discriminating between early and late effects (for overviews, see Conklin & Pellicer-Sánchez, 2016 and Pellicer-Sánchez & Siyanova-Chanturia, 2018). For example, Morales et al. (2016) examined how the grammatical gender of L1 modulates language processing during auditory comprehension of L2. They found that as early as 360 milliseconds (ms) after the article onset, bilingual participants dedicated more fixations to gender-congruent versus gender-incongruent targets. These findings suggest evidence for the coactivation of grammatical gender in L1 during auditory comprehension in L2.

Previous studies have used different methods to calculate cognate PO . For example, some studies asked participants to rate the phonological similarity of cognates (Dijkstra et al., 1999, 2010; Schwartz et al., 2007), while others employ an expert phonetician to calculate the overlap (Comesaña et al., 2012, 2015; Valente et al., 2018). The PO between targets and competitors has also been calculated based on shared phonemes of the word onset and then measured in milliseconds in the actual audio files (Blumenfeld & Marian, 2007). In the present study, we calculated the PO based on the translations' transcription and by precisely matching phonemes to the same positions. Midgley et al. (2011) used a similar procedure to calculate the orthographic overlap. We used this restrictive method because of the literature surrounding bilinguals' sensitivity to fine-grained phonemic variations (Ju & Luce, 2004; Weber & Cutler, 2004). This method was also adopted because our language combination was English and Spanish, where English has at least 10 more vowels than Spanish (Ladefoged, 2005, in Macizo et al., 2012). Thus, it was necessary to control for these variations and differences.

Based on previous studies, we expected shorter response times for words with high-PO compared to low-PO and no-PO linearly. We also expected proficiency to modulate this pattern; based on the BIA-d model predictions (Grainger et al., 2010) we anticipated attenuation of the CFE, such that the highly proficient group would exhibit smaller differences between high-PO, low-PO, and no-PO words. In contrast, the low proficient group would show considerable differences between the three conditions.

In terms of ET data, we looked at the proportion of fixations on the target pictures as done in previous studies (e.g., Morales et al., 2016) to have a measure of lexical access in time. Shorter fixations have been associated with cognates, suggesting less cognitive effort in comparison with noncognates. This pattern has been reported in natural reading (Cop et al., 2017), and self-paced reading (Bultena et al., 2014). In contrast to these studies, we expected a higher proportion of fixations to high-PO conditions, followed by low-PO, and no-PO conditions, respectively. We predicted this pattern because our study design was similar to Morales et al. (2016), in which a higher proportion of fixations were related to the less demanding condition. In the visual world paradigm, the faster the competition between targets and distractors is resolved, the faster the participants can select the target picture with a mouse, and this action is naturally accompanied by fixations that guide the selection. Thus, the quicker the target is recognized, the higher the proportion of fixations there are at that time. Similarly, in terms of response times (RTs), we expect that this effect would be further modulated by proficiency, with attenuated differences for highly proficient bilinguals and more considerable differences for low proficient bilinguals. In our monolingual control group, we expected no differences between conditions either in response times or in the proportion of fixations.

Method

Participants

Fifty-two native Spanish speakers from the University of Granada participated in the study in exchange for partial course credit or a small monetary reward. Four participants were excluded from the study because English was not their primary L2 or their data failed to record. Of the remaining 48 participants, 13 were male, and their ages ranged from 19 to 32 years old (mean (M) = 22.40, standard deviation (SD) = 2.81). All participants were native speakers of Spanish (L1), and English was their primary L2. All participants had normal or corrected to normal vision, and all (except two) were right-handed.

Their proficiency in L2 was assessed by a Michigan English Language Institute College English Test (MELICET) and verbal fluency task. A language questionnaire was also administered in which participants rated their auditory comprehension, writing, and speaking in the L2. The questionnaire consisted of a 10-point Likert scale where a score of 10 signified native proficiency (Language Experience and Proficiency Questionnaire (LEAP-Q); Marian et al., 2007). We calculated a composite proficiency score used in previous studies to split the participants into low and highly proficient groups (McMurray et al., 2010; Pivneva et al., 2012). For the composite score, we calculated a z value from self-assessment measures, MELICET, and English verbal fluency and added the results into a single score. Based on this score, we then median-split the group.

Moreover, 32 native English speakers (five males) were recruited from Pennsylvania State University, USA, to form a control group. Their ages ranged from 19 to 27 years old (M = 21.31, SD = 2.42). All participants had normal or corrected to normal vision, and all (except one) were right-handed. We administered the language background questionnaire (LEAP-Q; Marian et al., 2007) and a verbal fluency task in English to this group. The language proficiency measures and demographic variables for all three groups are displayed in Table 1.

Based on our median split, we reached two groups of bilinguals that differed significantly on MELICET score, $t(47) = 11.64$, standard error (SE) = 1.041, $p < 0.001$, verbal fluency task in English, $t(47) = 5.578$, $SE = 0.906$, $p < 0.001$, and all self-assessed measures of proficiency in English (all ps < 0.001). As expected, highly proficient bilinguals were more exposed to English than low proficient bilinguals, $t(47) = 2.395$, $SE = 2.107$, $p = 0.049$.

Table 1. Mean scores (with the standard deviation in parentheses) for language and demographic variables of low and highly proficient bilinguals and a monolingual control group.

	Low proficient bilinguals ($n=24$)	Highly proficient bilinguals ($n=24$)	Monolingual control group ($n=32$)
Age	22.00 (2.81)	22.79 (2.83)	21.31 (2.42)
MELICET score	26.46 (4.93)***	38.58 (4.39)	–
Verbal fluency in English	11.72 (2.49)***†††	16.77 (2.80)††	19.68 (3.81)
Self-assessed capacity			
– to speak in English	5.64 (1.71)***†††	7.98 (0.91)†††	9.79 (0.49)
– to understand spoken English	6.63 (1.71)***†††	8.38 (0.97)†††	9.86 (0.35)
– to write in English	7.08 (1.36)***†††	8.60 (0.82)†††	9.79 (0.41)
Exposure to English (%)	14.87 (7.44)* †††	19.92 (10.48)†††	99.48 (0.18)

Notes: MELICET = a multiple-choice English examination with a maximum score of 50 points. Comparisons of low proficient bilinguals with highly proficient bilinguals: *$p < 0.05$, ***$p < 0.001$; comparisons of low and highly proficient bilinguals with monolingual control group: ††$p < 0.01$, †††$p < 0.001$.

In comparison to our monolingual English speakers, both groups of bilinguals differed significantly in verbal fluency task in English, $t(55) = 9.394$, $SE = 0.847$, $p < 0.001$; $t(55) = 3.431$, $SE = 0.847$, $p = 0.003$ (for low and highly proficient bilinguals, respectively), and differed significantly on all self-assessed measures of proficiency in English and exposure to English (all ps < 0.001). These comparisons show that our bilingual groups were mostly exposed to L1 and did not reach a level of L2 comparable to native speakers.

Material

We selected pictures in grey-scale from the Multilingual Picture database (Duñabeitia et al., 2017; the complete set with its norms in English and Spanish is available at https://www.bcbl.eu/databases/multipic). We used the CLEARPOND Internet-based tool to acquire phonological information for the complete set of words (Marian et al., 2012). We adapted a formula used in a previous study to calculate the orthographic overlap and the phonological overlap between English and Spanish translations (Midgley et al., 2011). As such, we calculated the number of shared phonemes at the exact position between translation equivalents divided by their average phonological length, and we multiplied it by 100 to obtain the PO in percentage. We then created three experimental conditions with 20 words per condition: (a) high-PO condition with an overlap of 63.51%; (b) low-PO condition with an overlap of 37.42%; and (c) two conditions with no-PO overlap between English and Spanish translations (one was created as a distractor condition that was the same for all three conditions). A one-way analysis of variance (ANOVA) showed that PO differed significantly across conditions: $F(3, 76) = 553.0$, $p < 0.001$. *Post-hoc t*-tests revealed significant differences between all pairs of conditions (all ps < 0.001) except between the two conditions with no PO, $t(39) = 6.47 * 10^{-14}$, $p = 1.00$, as expected. Additionally, we controlled for orthographic similarity (OS; Van Orden, 1987) and normalized Levenshtein distance (NLD, Levenshtein, 1966). OS and NLD were computed using the online tool NIM (Guasch et al., 2013). As expected, the one-way ANOVAs performed on both OS and NLD showed significant differences between conditions, $F(3, 76) = 141.2$, $p < 0.001$; $F(3, 76) = 87.00$, $p < 0.001$, respectively. *Post-hoc t*-tests revealed significant differences between: high-PO and no-PO conditions, $t(39) = 15.141$, $p < 0.001$, $t(39) = 12.301$, $p < 0.001$, for OS and NLD, respectively; low-PO and no-PO conditions, $t(39) = 13.291$, $p < 0.001$, $t(39) = 10.084$, $p < 0.001$, for OS and NLD, respectively; no significant differences were found between the two no-PO conditions, $t(39) = 0.546$, $p = 0.947$, $t(39) = 0.244$, $p = 0.995$; and, interestingly, no significant differences were found between high-PO and low-PO conditions, $t(39) = 1.851$, $p = 0.258$, $t(39) = 2.217$, $p = 0.128$, for OS and NLD, respectively. These results demonstrate that high-PO and low-PO conditions differ phonologically but not orthographically, which is critical for the present study.

We controlled the experimental conditions in the following six psycholinguistic variables in English and Spanish: the age of acquisition; logarithmic frequency; relative frequency; orthographic length; phonological length; and phonological neighbourhood density. For the Spanish age of acquisition (AoA), we used a database by Alonso et al. (2015). For the English AoA, we used the database by Kuperman et al. (2012). We obtained the logarithmic frequency, relative frequency, and orthographic length with an online tool NIM (Guasch et al., 2013). This tool uses the LEXESP database for Spanish lexical frequencies (Sebastián-Gallés et al., 2000) and the British National Consortium database for English lexical frequencies (British National Consortium, 2007). Finally, we obtained both the phonological length and neighbourhood density with the online tool CLEARPOND (Marian et al., 2012). We also controlled the duration of the audio files used in the experiment. The one-way ANOVA on each one of these variables showed no significant differences across conditions (all ps > 0.22). The mean scores and respective standard deviation for each variable per condition are displayed in Appendix 1.

The visual material across all conditions was controlled for seven variables available in the MultiPic database (H-index, the percentage of modal names, valid responses, different responses, unknown responses, idiosyncratic responses, and visual complexity). The one-way ANOVAs conducted on these variables showed no significant differences across conditions (all ps > 0.14). The mean scores (with the standard deviation in parentheses) for these variables can be seen in Appendix 2.

Each experimental condition was composed of 20 different picture pairs of the target–distractor. While the target pictures varied among conditions (high-PO, low-PO, and no-PO), the distractors were the same across the experimental conditions so that any between-condition effect could be attributable to the presence of different targets. Moreover, the target and distractor pairs never shared an initial phoneme and were not related semantically. An example of the experimental material can be found in Figure 1.

The distractor images were repeated to prevent anticipation across the three experimental conditions. We created an additional filler condition where the distractors served as a target (filler condition A). We created two additional filler conditions (filler conditions B and C) where target words and distractors shared approximately the same amount of phonological information (43% and 41%, respectively). We utilized this method to mask the study's hypothesis from participants and to have more stimuli that would make any anticipation even more difficult. In total, we presented 120 pairs of pictures and the complete set of stimuli which can be found in Appendix 3. Moreover, the position of the drawings on the screen was counterbalanced, and each trial was presented at random. The pictures had dimensions of 250×250 pixels and a distance of 285 pixels in between, and they were presented on a 23", 1920×1080 pixels widescreen monitor.

The auditory material was recorded with a female native speaker of American English in an isolated room in a single session with a stereo recorder at the quality of 48 kHz.

Procedure

First, we asked the participants to fill in the language background questionnaire LEAP-Q (Marian et al., 2007) and MELICET English examination before the session in the laboratory to control for their proficiency in English and possible knowledge about any third language. After reading and signing an informed consent in the laboratory, participants completed the visual world experiment

Figure 1. An example of an experimental trial. The participants listened to the words 'telescope' while they had to select the appropriate image with a mouse. **Telescope**, *Telescopio* (a cognate word with high phonological overlap); and **Breakfast**, *Desayuno* (distractor with no phonological overlap).

in a dimmed isolated room. The instruction language was English. The participants sat approximately 65 cm from the screen, and their chin and forehead were stabilized to prevent any head movement. Participants completed a familiarization block that consisted of 10 trials that did not appear in the following task. We then calibrated the camera, and then the participants completed the task. Eye movements were recorded from the right eye using EyeLink 1000 (SR Research, Ontario, Canada). The right eye was selected since it is the dominant eye in the majority of people.

Moreover, the experimenter drift corrected the calibration before each trial. Each trial started with a fixation point, then two pictures appeared on the screen during 500 ms for a preview, and the target word in English was presented (the same interval was used in Morales et al., 2016). Using a mouse, the participants had to select an image that corresponded to the target word as fast and as accurately as possible. The pictures stayed on the screen until one of them was selected or until 4000 ms passed. The eye-fixations were recorded automatically by the software and were classified into three groups: target area; distractor area; and outside area. After the visual word task, the participants completed the verbal fluency task. By the end of the session, participants were briefed and received a small monetary reward or a partial course credit as compensation for their time. The laboratory session lasted from 40 to 50 minutes.

Results

Accuracy and RTs data analyses

We used R version 3.6 (R Core Team, 2019) and R Studio (RStudio Team, 2016) for all data analyses. We used the "lme4" package (Bates et al., 2015) to analyse accuracy and RTs with mixed-effects models. We used mixed-effects models to account for repeated measures design and to address non-independent data points per participants and items. For the accuracy analysis, we fitted a generalized linear mixed-effects model because the dependent variable was binomial (i.e., "correct" and "incorrect" responses). For the RTs analysis, we fitted a linear mixed-effects model because the dependent variable was continuous. In each model, the fixed effect predictors were Proficiency (with three levels: English monolinguals, highly proficient bilinguals, and low proficient bilinguals), and Cognate Type (also with three levels: high-PO, low-PO, and no-PO).

For the random effect component, we included random intercepts for participants and items (i.e., the target word).

To analyse main effects, we performed a likelihood ratio test (by calling "anova()" function) on a model that contained the fixed effect in question (Proficiency or Cognate Type), against a model that did not include it. Similarly, to analyse the Proficiency*Cognate Type interaction, we compared the full model with the interaction term to a model that contained only the main effects (see Winter, 2013, 2019). For *post-hoc* comparisons, we used the "emmeans" package (Russell, 2020) with Tukey's multiple comparison correction. In the case of significant interactions, we made planned comparisons of Cognate Type aggregated by our experimental groups (Proficiency).

Accuracy and RTs results

The accuracy for English monolinguals was 99.94%, for highly proficient bilinguals 99.51%, and for low proficient bilinguals 97.78%. The analysis described above revealed a main effect of Proficiency, Chi-squared test (χ^2) (2) = 30.972, $p < 0.001$. *Post-hoc t*-tests revealed that low proficient bilinguals were more likely to make errors than both highly proficient bilinguals (logit estimate: 1.60, $SE = 0.507$, $p = 0.005$) and English monolinguals (logit estimate: 3.81, $SE = 1.044$, $p < 0.001$). No significant differences were found between highly proficient bilinguals and English monolinguals in the likelihood of making errors (logit estimate: 2.21, $SE = 1.097$, $p = 0.108$) as both groups were very accurate. There was no main effect of Cognate Type χ^2 (2) = 0.888, $p = 0.641$, and no Proficiency*Cognate Type interaction χ^2 (4) = 4.2431, $p = 0.374$.

For the RTs analysis, RTs slower than 2.5 standard deviations of the mean per participant per condition were considered outliers and were excluded (2.01% of the data). The analysis (as described above) yielded a main effect of Proficiency, $\chi^2(2) = 24.653$, $p < 0.001$, no effect of Cognate Type, $\chi^2(2) = 1.299$, $p = 0.522$, and crucially, a significant Proficiency*Cognate Type interaction, $\chi^2(4) = 39.938$, $p < 0.001$. *Post-hoc t*-tests on Proficiency revealed that low proficient bilinguals were overall slower in responding than both highly proficient bilinguals, $t(80) = 4.090$, $SE = 38.7$, $p < 0.001$, and English monolinguals, $t(80) = 5.132$, $SE = 36.2$, $p < 0.001$. No statistically significant differences were found between highly proficient bilinguals and English monolinguals, $t(79.8) = 0.760$, $SE = 36.2$, $p = 0.728$. *Post-hoc t*-tests on the Proficiency*Cognate Type interaction further revealed a CFE in low proficient bilinguals as this group was faster in responding to both high-PO and low-PO words as compared to no-PO words, $t(87.8) = 3.199$, $SE = 28.7$, $p = 0.005$; and $t(87.2) = 3.100$, $SE = 28.7$, $p = 0.007$, respectively. And they were equally fast in responding to high-PO, and low-PO words, $t(87.2) = 0.104$, $SE = 28.7$, $p = 0.994$, therefore this CFE was not modulated by the degree of PO in RTs in this group. Surprisingly, highly proficient bilinguals did not significantly differ in any condition (all ps > 0.786), thus showed no CFE. As expected, English monolinguals did not significantly differ in any condition (all ps > 0.957). The mean RTs with standard deviations for each group and condition can be seen in Table 2.

We also performed an analysis with Proficiency as a continuous variable. In this analysis, we only included highly proficient bilinguals and low proficient bilinguals. The analysis did not differ from the previous analysis as it revealed a main effect of Proficiency, $\chi^2(2) = 8.782$, $p = 0.003$. This showed that as proficiency increases, the RTs in general decrease ($t(47.97) = -3.104$, $\beta = -69.22$, $SE = 22.30$, $p = 0.003$). Similarly, there was no main effect of Cognate Type, $\chi^2(2) = 2.361$, $p = 0.307$, but again, this analysis revealed a significant Proficiency*Cognate Type interaction, $\chi^2(4) = 17.015$, $p < 0.001$. The interaction can be seen in Figure 2; it reveals that as proficiency increases, the differences between no-PO condition and both high-PO and low-PO conditions decrease. On the other hand, we can again see that the differences between high-PO and low-PO conditions are negligible, as in the previous analysis.

Table 2. Mean response times in milliseconds (with the standard deviation in parentheses) for each group and condition.

	High-phonological overlap (PO) condition	Low-PO condition	No-PO condition	Mean (group)
Low proficient bilinguals	1087.9 (296.2)	1092.3 (334.7)	1178 (380.0)	1118.86 (340.78)
Highly proficient bilinguals	968.9 (257.5)	952.5 (269.4)	964.9 (263.3)	962.34 (263.07)
English monolinguals	941.8 (199.4)	939.1 (202.1)	932.2 (207.8)	934.56 (199.06)

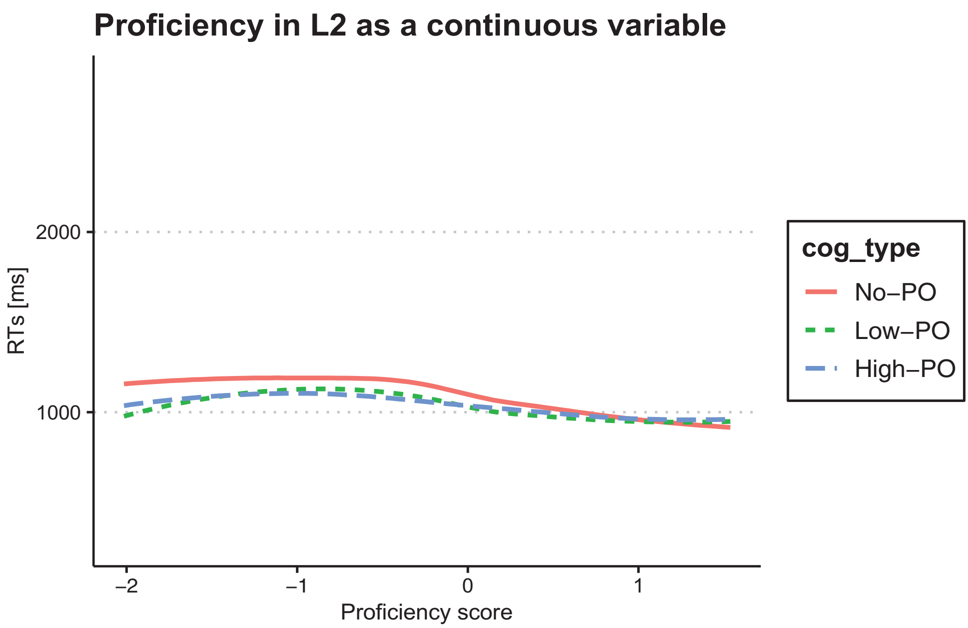

Figure 2. Response times for each condition as a function of proficiency in a second language as a continuous variable.

ET data analyses

Time-course of the proportion of fixations data analysis. For ET data analysis, the fixations were recorded automatically by SR Research software at the sampling rate of 1000 Hz. As such, we obtained one fixation per millisecond. Target pictures fixations were coded as 1, distractors fixations were coded as 2, and fixations to an area outside the target areas were coded as 0. All coding was included in the data analysis. Blinks, saccades, and error responses were excluded from the data analysis. We used "eyetrackingR" package (Dink & Ferguson, 2018) to analyse the exact time where conditions differed significantly from each other by performing a bootstrapped cluster-based permutation analysis (Maris & Oostenveld, 2007). This analysis offers an excellent compromise between type I error and sensitivity; it is carried out in two steps. In a first step, this analysis compares conditions by running *t*-tests for each time bin and merges statistically significant adjacent time bins into clusters. In a second step, it creates a Monte Carlo distribution by randomly shuffling time bins and repeating the first step over this newly generated data; this is done many times (for our data: 10 000 times). Finally, the t-values of our original clusters are compared to the distribution, which results in a Monte Carlo p-value that reflects the probability of observing our clusters solely by chance (also described in Poltrock et al., 2018).

As a dependent measure, we looked at the proportion of fixations within each 25 ms bin from the onset of the spoken stimuli until 1500 ms, reaching 60-time bins. We established a t-threshold with an alpha level of 0.05 (two-tailed). The analysis revealed that for low proficient bilinguals, the low-PO and no-PO conditions diverged from 700 to 925 ms (cluster $t=25.996$, Monte Carlo $p=0.032$). For highly proficient bilinguals, the low-PO and no-PO conditions diverged from 175 to 400 ms (cluster $t=24.743$, Monte Carlo $p=0.041$). And surprisingly, for English monolinguals, both high-PO and low-PO conditions diverged from the no-PO condition from 25 to 250 ms (cluster $t=23.103$, Monte Carlo $p=0.039$, and cluster $t=22.051$, Monte Carlo $p=0.041$, respectively). For each group, the proportion of fixations in time can be seen in Figures 3 to 5.

Latency in shifting analysis. We also carried out an analysis of the latency in shifting from the distractor to the target image. This measure gives us an estimate of a time when target words have been recognized and could discriminate early phonological coactivation. To carry out this analysis, we used "eyetrackingR" package (Dink & Ferguson, 2018) in R. First, we filtered trials in which participants had been looking at distractor images at the onset of the spoken word and then switched to target pictures. From these trials, we excluded those in which participants shifted to the target image before 200 ms as these switches were probably programmed before the spoken word could be processed (see Poltrock et al., 2018, for similar analysis). These distractor-initial trials corresponded to 22.98% of all the trials (22.08% for low proficient bilinguals, 20.55% for highly proficient bilinguals, and 22.86% for English monolinguals). For each group, the mean shifting latencies with standard deviations can be seen in Table 3.

To analyse the latency in shifting from distractor to target images, we fitted a linear mixed-effects models on the latency data with the same fixed and random effects structure as for the accuracy and RTs analyses described above. The analysis showed a main effect of Proficiency, $\chi^2(2)=10.74$, $p=0.005$. *Post-hoc* t-tests revealed that low proficient bilinguals were slower than English monolinguals, $t(68.1)=3.316$, $SE=15.8$, $p=0.004$, and marginally slower than highly proficient bilinguals, $t(72.4)=3.316$, $SE=17.0$, $p=0.059$. There were no statistical differences between highly proficient bilinguals and English monolinguals, $t(71.8)=0.822$, $SE=13.1$, $p=0.690$. There was neither no main effect of Cognate Type, $\chi^2(2)=0.430$, $p=0.806$, nor significant Proficiency*Cognate Type interaction, $\chi^2(4)=1.988$, $p=0.738$.

Discussion

The present study aimed to explore the CFE during L2 auditory comprehension in bilinguals. We also examined whether proficiency in L2 and the degree of PO modulated a possible CFE. Importantly, we designed a visual world paradigm task that allowed us to explore the time-course of this effect by measuring the proportion of fixations in time.

Both RTs and the ET data (latencies in shifting from distractor-initial trials to targets) showed that low proficient bilinguals were overall slower in identifying words in L2 and made more errors than highly proficient bilinguals and English monolinguals. These results are in line with the L2 models of language processing (BIA-d, Grainger et al., 2010; RHM, Kroll & Stewart, 1994; Kroll et al., 2010) assuming that at early stages of L2 acquisition, access to the meaning of the words depends to a great extent on L1 word-form mediation. Following this reasoning, the overall slower responses of low proficient bilinguals in the present study suggest that they relied on such L1-mediated translation.

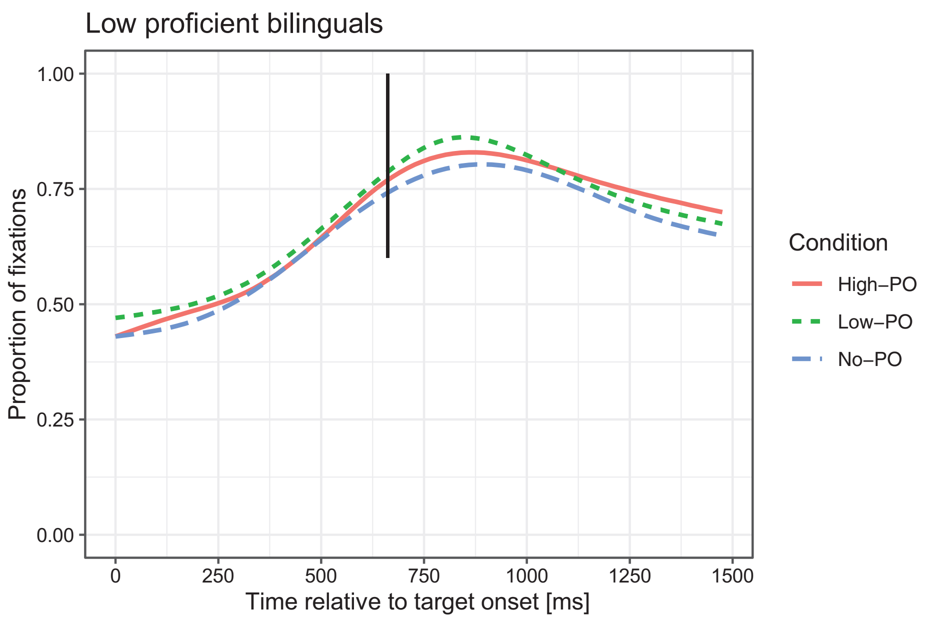

Figure 3. The proportion of fixations as a function of time for low proficient bilinguals and each experimental target. The green-shaded area marks significant differences between low-PO and no-PO conditions. The vertical line marks the mean word offset and the grey-shaded area its standard deviation.
Note: high-PO, high phonological overlap; low-PO, low phonological overlap; and no-PO, no phonological overlap.

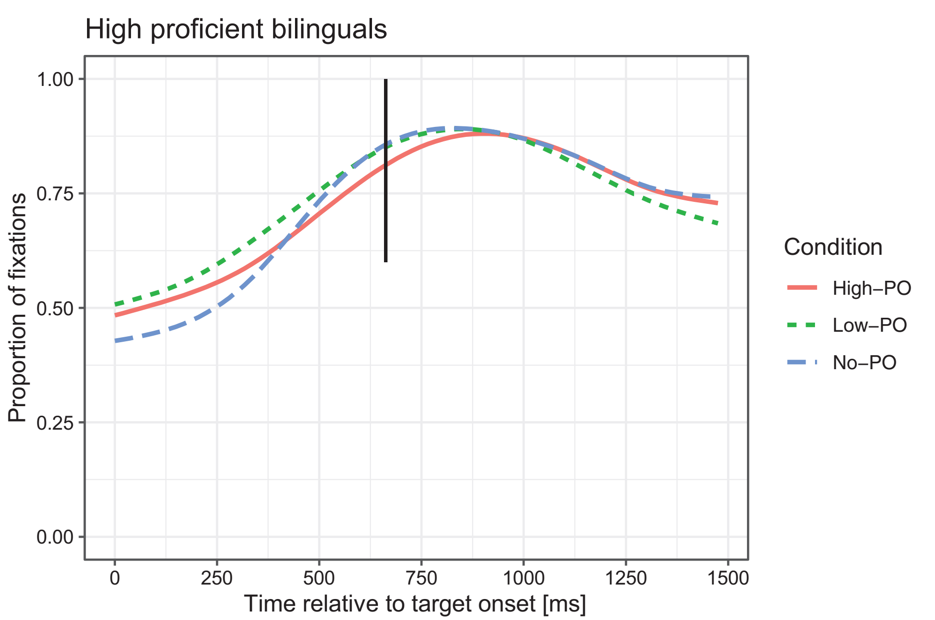

Figure 4. The proportion of fixations as a function of time for low proficient bilinguals and each experimental target. The green-shaded area marks significant differences between low-PO and no-PO conditions. The vertical line marks the mean word offset and the grey-shaded area its standard deviation.
Note: high-PO, high phonological overlap; low-PO, low phonological overlap; and no-PO, no phonological overlap.

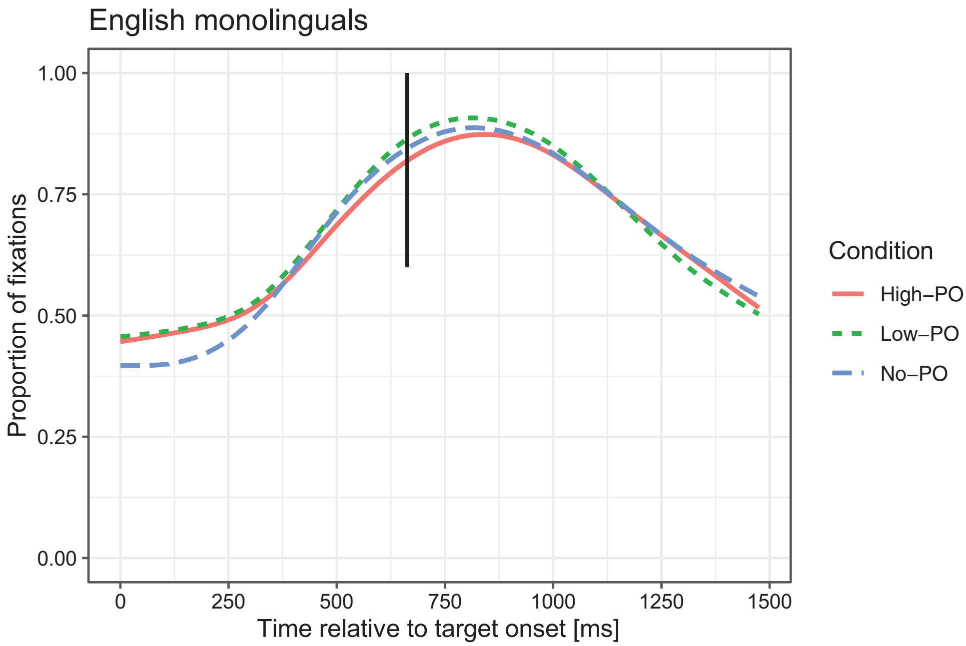

Figure 5. The proportion of fixations as a function of time for low proficient bilinguals and each experimental target. The violet-shaded area marks significant differences between both high-PO and low-PO conditions relative to no-PO condition. The vertical line marks the mean word offset and the grey-shaded area its standard deviation.
Note: high-PO, high phonological overlap; low-PO, low phonological overlap; and no-PO, no phonological overlap.

Table 3. Mean latencies in shifting from distractor-initial trials to target pictures in milliseconds (with the standard deviation in parentheses) for each group and condition.

	High-phonological overlap (PO) condition	Low-PO condition	No-PO condition	Mean (group)
Low proficient bilinguals	415.03 (175.43)	418.83 (198.35)	408.43 (176.01)	413.81 (187.07)
Highly proficient bilinguals	387.62 (149.56)	369.21 (149.45)	379.23 (135.37)	378.65 (143.66)
English monolinguals	359.12 (117.99)	363.56 (126.13)	372.18 (137.23)	365.27 (127.62)

Crucially, low proficient bilinguals showed a clear CFE as this group was faster in responding to cognates as revealed by RTs. While inspecting the cognate facilitation time-course, we found a robust effect between noncognates and cognates with low PO. This effect lasted from 700 ms to 925 ms post-stimulus presentation. The lateness of the effect converged with the latencies in shifting from distractor-initial trials that we carried out as an additional measure of early phonological coactivation, and it showed no differences between conditions.

Furthermore, the degree of PO seemed not to determine the amount of facilitation, at least at the behavioural level. Low proficient bilinguals were equally fast in identifying words with high and low PO. These behavioural data are in line with Valente et al. (2018) who found cognate facilitation for both adults and children during cognate acquisition but adults in their study were not sensitive to the degree of phonological and orthographic overlap in cognates. In our study, the low proficient bilinguals, all adults, showed the same pattern in RTs. Taking into account the ET data in low proficient bilinguals, the higher proportion of fixations for low-PO condition further reflects cognate facilitation and its origin in the late time-window as described above. However, the absence of a similar effect between high-PO condition and noncognates is surprising since we predicted that this condition would have been the easiest to identify and thus would receive the highest proportion of fixations. One may argue that since we derived the formula to calculate the PO from a study that originally used it to calculate orthographic overlap (Midgley et al., 2011), it would be possible that the PO in our high-PO condition could have been, in fact, lower in the initial proportion of the target words in comparison to low-PO condition. This could explain the observed differences in the proportion of fixations since the initial part of the word is crucial during auditory comprehension (Marslen-Wilson, 1987). However, we verified that this was not the case as the proportion of shared phonemes at each position was always higher in the high-PO condition than in the low-PO condition. Although the difference in the proportion of fixations proved to be an ideal measure to capture prolonged effects (e.g., Morales et al., 2016), it may not easily capture fine-grained effects. For example, Blumenfeld and Marian (2007) reported that the amount of PO was positively related to language coactivation when considering prolonged time-windows. Still, no coactivation of L2 was found in low and medium overlapping conditions in German native speakers when analysing the proportion of fixations in time. Previous studies have also shown significant results in one ET measure in the absence of significant findings in a closely related measure, suggesting that they differ in their sensibility to capture the same underlying process. For example, Cop et al. (2017) studied cognate facilitation in L1 and L2 naturalistic reading. They found that when participants read in L2, the amount of orthographic overlap was related to the first fixation duration or the more orthographic overlap, the shorter the fixations. However, it was not related to single fixation duration, nor gaze duration. In contrast, when looking at the probability of skipping, only the words with high overlap that were shorter than six characters were skipped more often. In other words, the amount of orthographic overlap was related to only two out of the four measures tapping into early automatic processes. Perhaps other techniques and procedures such as electroencephalogram (EEG) may be more suitable to explore the time-course of cognate facilitation. Despite the fact that the ET data are not entirely consistent with our predictions, and that other measures could offer a more precise picture of the time course of the CFEs, the analyses carried out in this study suggest that the cognate facilitation appeared at a late stage of processing in low proficient bilinguals.

In contrast, highly proficient bilinguals did not show cognate facilitation in any measure and their performance did not differ from English monolinguals. Although we expected that high proficient bilinguals would show cognate facilitation, we could speculate that the absence of this effect is not at odds with the idea that when fluency in L2 is acquired, lexical–conceptual links in L2 are established, and meaning is directly accessed (Grainger et al., 2010; Kroll & Stewart, 1994; Kroll et al., 2010). What we did not predict was that the highly proficient group would have reached this level of proficiency because they still differed on every L2 proficiency measure compared to the English monolinguals, and their exposure to English was still quite low (around 20%). While Grainger et al. (2010) speculated that immersion in L2 might be necessary to reach the 'magical point' when L2 production and comprehension becomes effortless, we believe that this may not be the case for L2 comprehension. We should still keep in mind that the materials used in this study were composed of concrete and simple nouns. Thus, any generalizations made about the absence of cognate facilitation during auditory comprehension of L2 in highly proficient bilinguals must be taken with caution, and more challenging material should be used in future studies to rule out possible ceiling effects. Therefore, our findings are at odds with studies that have already reported the coactivation of L1 with highly proficient bilinguals during auditory comprehension (Marian & Spivey, 2003a, 2003b; Weber & Cutler, 2004) and are in line with studies that have found no language coactivation of L1 (Marian & Spivey, 2003a).

Surprisingly, we have observed early differences in the proportion of fixations for cognate and noncognate words in English monolinguals and highly proficient bilinguals. The differences spanned from 25 ms to 250 ms and 175 ms to 400 ms post-target presentation for each group, respectively. In theory, we should not have observed any differences in ET data during the first 200 ms post-target presentation since programming a saccade takes on average that much time (Carpenter, 1988). While in English monolinguals, this effect does not last long and vanishes quickly after the 200 ms mark, in highly proficient bilinguals, it spans until 400 ms. We are aware that this is an issue since baseline differences could be masking putative early effects. However, similar to our RTs findings, we have observed no differences between conditions in our second measure of early effects (the latencies in shifting from distractor-initial trials) for highly proficient bilinguals nor for the English monolinguals. Nevertheless, the possibility of early effects still exists especially in the light of previous studies that have already reported early cognate effects; in particular, in the P200 event-related potential (ERP) component in naming (Strijkers et al., 2010). On the other hand, Comesaña et al. (2012) found an interaction, even though marginal, between orthography and phonology in the early N100 ERP components in a silent reading task with masked priming. Therefore, other techniques such as EEG could be more suitable to evaluate the time-course of cognate facilitation in bilinguals during auditory comprehension, especially for detecting early effects that are not always reflected in RTs.

Recently, a proposal of a new bilingual computational model Multilink (Dijkstra et al., 2019) was presented. It integrates the two models of visual word recognition (BIA/BIA+) and oral production (RHM). It is noteworthy that auditory bilingual comprehension is not considered in the Multilink model yet. Therefore, it would be exciting to see an all-encompassing model of bilingual production and comprehension at all levels, with the inclusion of bilingual writing production (see for example, Iniesta et al., under review).

In conclusion, we have shown a CFE during auditory comprehension of L2 that was modulated by proficiency as only low proficient bilinguals showed late cognate facilitation. The degree of PO seemed not to play a role in the access to cognates, at least at the behavioural level. The finding of no cognate facilitation in highly proficient bilinguals should motivate future studies to employ other techniques and procedures along with more difficult material during bilingual auditory comprehension and explore whether this manipulation could show cognate facilitation due to shared phonology.

Acknowledgements

We thank McKenzie Himelein-Wachowiak for her help with the data acquisition, Brooke E. Leonard for recording the auditory stimuli used in this study, Daniel Caldera for assistance with the English, as well as Ad Backus, Montserrat Comesaña and an anonymous reviewer for very useful comments and suggestions on earlier versions of this manuscript.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by Grants PCIN-2015-165-C02-01, PCIN-2015-132 and PSI2017-89324-C2-1-P from the Spanish Ministry of Economy and Competitiveness.

ORCID iDs

Filip Andras https://orcid.org/0000-0002-8876-7817

Daniela Paolieri https://orcid.org/0000-0002-3964-7405

References

Alonso, M. A., Fernandez, A., & Díez, E. (2015). Subjective age-of-acquisition norms for 7,039 Spanish words. *Behavior Research Methods*, *47*(1), 268–274. https://doi.org/10.3758/s13428-014-0454-2

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Blumenfeld, H. K., & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes*, *22*(5), 633–660. https://doi.org/10.1080/01690960601000746

British National Consortium (2007). The British National Corpus, version 3 (BNC XML Edition). Distributed by Oxford University computing services on behalf of the British National Consortium. http://www.natcorp.ox.ac.ukon.

Bultena, S., Dijkstra, T., & Van Hell, J. G. (2014). Cognate effects in sentence context depend on word class, L2 proficiency, and task. *Quarterly Journal of Experimental Psychology*, *67*(6), 1214–1241.

Carpenter, R. H. (1988). *Movements of the eyes* (2nd Rev.). Pion Limited.

Conklin, K., & Pellicer-Sánchez, A. (2016). Using eye-tracking in applied linguistics and second language research. *Second Language Research*, *32*(3), 453–467. https://doi.org/10.1177/0267658316637401

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

Cop, U., Dirix, N., Van Assche, E., Drieghe, D., & Duyck, W. (2017). Reading a book in one or two languages? An eye movement study of cognate facilitation in L1 and L2 reading. *Bilingualism: Language and Cognition*, *20*(4), 747–769. https://doi.org/10.1017/S1366728916000213

Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(5), 1283–1296. https://doi.org/10.1037//0278-7393.26.5.1283

Dijkstra, T., & Van Heuven, W. J. (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

Dijkstra, T., Grainger, J., & Van Heuven, W. J. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, *41*, 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., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2019). Multilink: a computational model for bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, *22*(4), 657–679. https://doi.org/10.1017/s1366728918000287

Dink, J., & Ferguson, B. (2018). eyetrackingR_. R package version 0.1.*8*. http://www.eyetracking-R.com.

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2017). MultiPic: A standardised set of 750 drawings with norms for six European languages. *The Quarterly Journal of Experimental Psychology*, *71*(4), 808–816. https://doi.org/10.1080/17470218.2017.1310261

Grainger, J., Midgley, K., & Holcomb, P. J. (2010). Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). In M. Kail, & M. Hickmann (Eds.), *Language Acquisition Across Linguistic and Cognitive Systems* (pp. 267–283) John Benjamins. https://doi.org/10.1075/lald.52.18gra

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*(3), 765–771. https://doi.org/10.3758/s13428-012-0296-8

Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition*, *106*(1), 501–511. https://doi.org/10.1016/j.cognition.2007.02.001

Iniesta, A., Rossi, E., Bajo, T., & Paolieri, D. (under review). The orthographic and phonological similarity in bilingual writing: evidence from late bilinguals and heritage speakers.

Jared, D., Cormier, P., Levy, B. A., & Wade-Woolley, L. (2012). Cross-language activation of phonology in young bilingual readers. *Reading and Writing*, *25*(6), 1327–1343. https://doi.org/10.1007/s11145-011-9320-0

Ju, M., & Luce, P. A. (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science*, *15*(5), 314–318. https://doi.org/10.1111/j.0956-7976.2004.00675.x

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497–514. https://doi.org/10.1080/20445911.2013.799170

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., Bobb, S. C., & Hoshino, N. (2014). Two languages in mind: Bilingualism as a tool to investigate language, cognition, and the brain. *Current Directions in Psychological Science*, *23*(3), 159–163. https://doi.org/10.1177/0963721414528511

Kroll, J. F., Dussias, P. E., & Bajo, M. T. (2018). Language use across international contexts: Shaping the minds of L2 speakers. *Annual Review of Applied Linguistics*, *38*, 60–79. https://doi.org/10.1017/s0267190518000119 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: Language and Cognition*, *13*(3), 373–381. https://doi.org/10.1017/S136672891000009X

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Ladefoged, P. (2005). *Vowels and consonants: An introduction to the sounds of languages* (2nd ed.). Blackwell.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady*, *10*(8), 707–710.

Macizo, P., & Bajo, M. T. (2006). Reading for repetition and reading for translation: Do they involve the same processes? *Cognition*, *99*(1), 1–34. https://doi.org/10.1037/e537272012-414

Macizo, P., Van Petten, C., & O'Rourke, P. L. (2012). Semantic access to embedded words? Electrophysiological and behavioral evidence from Spanish and English. *Brain and Language*, *123*(2), 113–124. https://doi.org/10.1016/j.bandl.2012.07.010

Marian, V., & Spivey, M. (2003a). Competing activation in bilingual language processing: Within-and between-language competition. *Bilingualism: Language and Cognition*, *6*(2), 97–115. https://doi.org/10.1017/s1366728903001068

Marian, V., & Spivey, M. (2003b). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics*, *24*(2), 173–193. https://doi.org/10.1017/S0142716403000092

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PloS ONE*, *7*(8), e43230. https://doi.org/10.1371/journal.pone.0043230

Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, *50*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067)

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG-and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1–2), 71–102. https://doi.org/10.1016/0010-0277(87)90005-9

McMurray, B., Samelson, V. M., Lee, S. H., & Tomblin, J. B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cognitive Psychology*, *60*(1), 1–39. https://doi.org/10.1016/j.cogpsych.2009.06.003

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

Morales, L., Paolieri, D., Dussias, P. E., Kroff, J. R. V., Gerfen, C., & Bajo, M. T. (2016). The gender congruency effect during bilingual spoken-word recognition. *Bilingualism: Language and Cognition*, *19*(2), 294–310. https://doi.org/10.1017/s1366728915000176

Pellicer-Sánchez, A., & Siyanova-Chanturia, A. (2018). Eye movements in vocabulary research. *ITL-International Journal of Applied Linguistics*, *169*(1), 5–29. https://doi.org/10.1075/itl.00004.pel

Pivneva, I., Palmer, C., & Titone, D. (2012). Inhibitory control and L2 proficiency modulate bilingual language production: Evidence from spontaneous monologue and dialogue speech. *Frontiers in Psychology*, *3*, 57. https://doi.org/10.3389/fpsyg.2012.00057

Poltrock, S., Chen, H., Kwok, C., Cheung, H., & Nazzi, T. (2018). Adult learning of novel words in a non-native language: Consonants, vowels, and tones. *Frontiers in Psychology*, *9*, 1211. https://doi.org/10.3389/fpsyg.2018.01211

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing.

RStudio Team (2016). RStudio: Integrated development for R. RStudio, Inc.

Russell, L. (2020). emmeans: Estimated marginal means, aka least-squares means. R package version 1.4.*5*. https://CRAN.R-project.org/package=emmeans

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*(1), 106–129. https://doi.org/10.1080/01690960500463920

Sebastián-Gallés, N., Carreiras, M., Cuetos, F., & Martı, M. A. (2000). *Lexesp. Léxico informatizado del español* [Lexesp. Computerized lexicon of Spanish]. Publicacions i Edicions UB. [In Spanish.]

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science*, *10*(3), 281–284. https://doi.org/10.1111/1467-9280.00151

Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. *Cerebral Cortex*, *20*(4), 912–928. https://doi.org/10.1093/cercor/bhp153

Tanenhaus, M. K., Magnuson, J. S., Dahan, D., & Chambers, C. (2000). Eye movements and lexical access in spoken-language comprehension: Evaluating a linking hypothesis between fixations and linguistic processing. *Journal of Psycholinguistic Research*, *29*(6), 557–580. https://doi.org/10.1023/A:1026464108329

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634. https://doi.org/10.1037/e537272012-052

Valente, D., Ferré, P., Soares, A., Rato, A., & Comesaña, M. (2018). Does phonological overlap of cognate words modulate cognate acquisition and processing in developing and skilled readers? *Language Acquisition*, *25*(4), 438–453. https://doi.org/10.1080/10489223.2017.1395029

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory and Cognition*, *15*(3), 181–198. https://doi.org/10.3758/bf03197716

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1–25. https://doi.org/10.1016/S0749-596X(03)00105-0

Winter, B. (2013). Linear models and linear mixed effects models in R with linguistic applications. arXiv:1308.*5499*. http://arxiv.org/pdf/1308.5499.pdf

Winter, B. (2019). *Statistics for linguists: An introduction using R*. Routledge.

Author biographies

Filip Andras is a PhD student in the Department of Experimental Psychology (University of Granada, Spain). His research is focused on the relationship between language and mathematics, and language processing in bilinguals.

Marta Rivera is a PhD student in the Department of Experimental Psychology (University of Granada, Spain). Her research interests are focused on the individual differences that underlie language learning and other aspects of bilingualism such as language production and reading.

Teresa Bajo is a full professor in the Department of Experimental Psychology and a member of the Mind, Brain and Behavior Research Center (University of Granada, Spain). She is an expert on cognitive neuroscience with a special focus on language, memory, and bilingualism.

Paola E Dussias is a professor of Spanish, Linguistics and Psychology. The primary focus of her research concerns bilingual sentence processing, cognitive aspects of bilingualism, and language contact phenomena such as code-switching.

Daniela Paolieri is an associate professor in the Department of Experimental Psychology (University of Granada, Spain). She currently carries out research on cognitive psychology in the areas of psycholinguistics and memory, she has focused especially on grammatical gender processing in language production and comprehension, addressing this issue in both monolingual and bilingual contexts.

Appendix 1. Mean scores (with the standard deviation in parentheses) for linguistic variables for each experimental condition.

	High phonological overlap (PO) condition	Low PO condition	No PO condition
PO (%)	63.51 (11.32)	37.43 (3.35)	0.00 (0.00)
Orthographic similarity	0.75 (0.15)	0.67 (0.15)	0.12 (0.13)
Normalized Levenshtein distance	0.77 (0.17)	0.66 (0.17)	0.14 (0.19)
Logarithmic frequency			
in English	0.98 (0.44)	0.89 (0.57)	1.13 (0.70)
in Spanish	0.92 (0.43)	0.94 (0.58)	1.19 (0.55)
Relative frequency			
in English	14.91 (19.28)	16.80 (27.29)	36.65 (52.55)
in Spanish	11.76 (11.25)	23.16 (46.60)	30.92 (41.99)
Age of acquisition			
in English	6.11 (1.40)	6.07 (1.76)	5.33 (1.62)
in Spanish	5.23 (1.36)	5.29 (1.63)	4.40 (1.64)
Orthographic length			
in English	6.10 (1.90)	6.15 (1.57)	6.70 (1.69)
in Spanish	6.05 (1.88)	6.50 (1.64)	5.80 (1.51)
Phonological length			
in English	5.55 (1.67)	5.20 (1.77)	5.40 (1.14)
in Spanish	6.05 (1.79)	6.20 (1.40)	5.55 (1.47)
Phonological neighbourhood density			
in English	4.40 (6.75)	3.70 (5.41)	4.65 (6.21)
in Spanish	1.15 (2.70)	1.60 (3.12)	1.30 (2.76)
Audio files length [ms]	633 (100)	677 (106)	681 (99)

Appendix 2. Mean scores (with the standard deviation in parentheses) for visual material available in MultiPic database (Duñabeitia et al., 2017) for each experimental condition.

	High phonological overlap (PO) condition	Low PO condition	No PO condition
H-index			
in English	0.58 (0.59)	0.45 (0.62)	0.70 (0.62)
in Spanish	0.22 (0.34)	0.36 (0.55)	0.39 (0.58)

(Continued)

Appendix 2. (Continued)

	High phonological overlap (PO) condition	Low PO condition	No PO condition
Modal names (%)			
in English	86.74 (16.51)	88.06 (18.65)	81.46 (19.83)
in Spanish	96.09 (6.30)	90.52 (15.70)	90.73 (15.77)
Valid responses (%)			
in English	94.65 (5.64)	94.55 (11.13)	94.10 (10.07)
in Spanish	97.40 (2.76)	97.25 (4.60)	97.30 (3.81)
Number of different responses			
in English	2.90 (1.80)	2.25 (1.40)	2.85 (1.57)
in Spanish	1.90 (1.33)	2.00 (1.62)	2.25 (1.86)
Unknown responses (%)			
in English	3.40 (4.97)	3.55 (9.48)	3.70 (8.90)
in Spanish	1.45 (1.54)	1.50 (1.93)	1.35 (1.27)
Idiosyncratic responses (%)			
in English	1.95 (2.24)	1.90 (2.07)	2.20 (2.55)
in Spanish	1.45 (1.54)	1.50 (1.93)	1.35 (1.27)
Visual complexity			
in English	2.78 (0.44)	2.87 (0.56)	2.55 (0.47)
in Spanish	2.48 (0.45)	2.56 (0.56)	2.31 (0.50)

Note: H-index, a statistic that reflects the level of agreement across participants, the closer to the 0, the higher the agreement.

Appendix 3. The condition with high phonological overlap (PO) between English and Spanish translations.

English target	Spanish translation	English distractor	Spanish translation
Kiwi	*Kiwi*	Fist	*Puño*
Dentist	*Dentista*	Briefcase	*Maletín*
Plant	*Planta*	Rubber	*Goma*
Submarine	*Submarino*	Fireplace	*Chimenea*
Dolphin	*Delfín*	Moustache	*Bigote*
Fruit	*Fruta*	Cloud	*Nube*
Telescope	*Telescopio*	Breakfast	*Desayuno*
Balcony	*Balcón*	Pumpkin	*Calabaza*
Button	*Botón*	Rocket	*Cohete*
Lemon	*Limón*	Frog	*Rana*
Melon	*Melón*	Record	*Disco*
Taxi	*Taxi*	Eyebrow	*Ceja*
Microphone	*Micrófono*	Biscuit	*Galleta*
Broccoli	*Brócoli*	Corridor	*Pasillo*
Pipe	*Pipa*	Honey	*Miel*
Cape	*Capa*	Handle	*Mango*
Pilot	*Piloto*	Queen	*Reina*
Sack	*Saco*	Necklace	*Collar*
Mosquito	*Mosquito*	Prison	*Cárcel*
Cactus	*Cactus*	Squirrel	*Ardilla*

The condition with low PO between English and Spanish translations.

English target	Spanish translation	English distractor	Spanish translation
Koala	*Koala*	Record	*Disco*
Diamond	*Diamante*	Corridor	*Pasillo*
Policeman	*Policía*	Fireplace	*Chimenea*
Miner	*Minero*	Fist	*Puño*
Camel	*Camello*	Rubber	*Goma*
Dinosaur	*Dinosaurio*	Necklace	*Collar*
Bottle	*Botella*	Eyebrow	*Ceja*
Pistachio	*Pistacho*	Briefcase	*Maletín*
Trophy	*Trofeo*	Handle	*Mango*
Volcano	*Volcán*	Pumpkin	*Calabaza*
Coconut	*Coco*	Moustache	*Bigote*
Guitar	*Guitarra*	Cloud	*Nube*
Tattoo	*Tatuaje*	Honey	*Miel*
Tram	*Tranvía*	Frog	*Rana*
Castle	*Castillo*	Squirrel	*Ardilla*
Gorilla	*Gorila*	Prison	*Cárcel*
Sun	*Sol*	Queen	*Reina*
Camera	*Cámara*	Breakfast	*Desayuno*
Panther	*Pantera*	Rocket	*Cohete*
Pear	*Pera*	Biscuit	*Galleta*

The condition with no PO between English and Spanish translations.

English target	Spanish translation	English distractor	Spanish translation
Shoulder	*Hombro*	Biscuit	*Galleta*
Newspaper	*Periódico*	Fireplace	*Chimenea*
Pencil	*Lápiz*	Queen	*Reina*
Cupcake	*Pastel*	Briefcase	*Maletín*
Spider	*Araña*	Record	*Disco*
Hairbrush	*Cepillo*	Necklace	*Collar*
Elbow	*Codo*	Honey	*Miel*
Computer	*Ordenador*	Breakfast	*Desayuno*
Wolf	*Lobo*	Cloud	*Nube*
Suitcase	*Maleta*	Prison	*Cárcel*
Milk	*Leche*	Rubber	*Goma*
Painting	*Cuadro*	Squirrel	*Ardilla*
Table	*Mesa*	Fist	*Puño*
Sunflower	*Girasol*	Moustache	*Bigote*
Glass	*Vaso*	Eyebrow	*Ceja*
Engine	*Motor*	Frog	*Rana*
Mattress	*Colchón*	Corridor	*Pasillo*
Coffin	*Ataúd*	Handle	*Mango*
Backpack	*Mochila*	Pumpkin	*Calabaza*
Zebra	*Cebra*	Rocket	*Cohete*