The role of cognateness in non-native spoken word recognition

Gonzalo Garcia-Castro
Center for Brain and Cognition
Universitat Pompeu Fabra
Barcelona, 08005
gonzalo.garciadecastro@upf.edu

Serene Siow
Department of Experimental Psychology
University of Oxford
Oxford, OX2 6GG
siow.serene@gmail.com

Kim Plunkett
Department of Experimental Psychology
University of Oxford
Oxford, OX2 6GG
kim.plunkett@psy.ox.ac.uk

Nuria Sebastian-Galles
Center for Brain and Cognition
Universitat Pompeu Fabra
Barcelona, 08005
nuria.sebastian@upf.edu

May 9, 2024

Abstract

Understanding spoken words in a non-native language is easier when the words are phonologically similar to their translations into the listener's native language are phonologically similar. Phonological similarity plays a more important role in the language performance of low-proficiency bilinguals, compared to that of high-proficiency bilinguals., and This suggests that low-proficiency bilinguals may points to the latter as relying more strongly on cross-language word-to-word links (lexical route), while high-proficiency bilinguals rely more on word-concept associations (conceptual route) than on cross-language word-to-word links (lexical route). Disentangling bilinguals' reliance on the conceptual and lexical and conceptual routes during translation is challenging, given that both sources of information are present during spoken word comprehensionto some extent in both high- and low-proficiency bilinguals. In this study, we tested English and Spanish native speakers in a translation elicitation task, in which they listened to words in an unfamiliar language (Spanish or Catalan for English natives, and Catalan for Spanish natives), and then had to translate them to their native language. Given their lack of previous knowledge ion the unfamiliar tested language, phonological similarity between the presented words and their correct translations was the only cue available for participants to succeed in the task. This allowed us to explore the informativeness of the lexical route for word translation, in the absence of information from the conceptual route. Our results indicate that participants benefited from the phonological similarity between translation equivalents: the more similar the word they listened to and its translational equivalent, the higher the probability of a correct translation. These results point to cross-language word-to-word links as providing rich information during word translation.

Keywords cognate • spoken word recognition • phonology • bilingualism • non-native speech processing • lexical access

1 Introduction

Listening to non-native speech is costlier than listening to native speech, even for highly proficient bilinguals, and especially in acoustically adverse situations (Lecumberri et al., 2010; Takata & Nábělek, 1990). However, there are some sources of information which may make non-native speech comprehension easier. One of the sources of this increased difficulty stems from the possible mismatch between the phonology of the native and

the non-native languages: some acoustic features embedded in the non-native speech signal do not overlap perfectly with any phonemic category in the listener's native language. For example, imagine a Spanish native, with no previous familiarity with any other language, who listens for the first time to French may encounter the word /p t/ (*porte*), translation of door in French. The voiced uvular fricative consonant / / and the open-mid back rounded vowel / / do not exist in Spanish as phonemic categories. These discrepancies can nonetheless be perceived as allophonic variations the native (Spanish) phonemes /r/ and /o/ (Best et al., 2001), and the non-native speech signal may engage lexical processing mechanisms (Weber & Cutler, 2004), leading to the activation and selection, and therefore the recognition of its translation to Spanish / pwer.ta/ (*puerta*). The recognition of this French, unfamiliar word relies almost entirely on its similarity with its Spanish translation. Conversely, imagine now that the same Spanish listener hears English for the first time and encounters the word /d / (*door*). This time, phonological similarity is of little help: / pwer.ta/ and /d / share very few phonemes. It would come as a surprise if the listener was able to translate the English word successfully, if not by relying on previously learnt word-concept associations (which they lack, as they are unfamiliar to English).

Cross-language phonological similarity seems, at face value, an important source of information for unfamiliar listeners during auditory word recognition, even if translations are not phonologically identical. Although this mismatch has a noticeable toll on comprehension (Cutler et al., 2004), the fact that word recognition can take place at all in such circumstances illustrates that non-native listeners–even low-proficiency ones–are rarely completely naïve to the language they are listening to. In the present study, we investigate the extent to which adult listeners rely on the phonological similarity between their native and non-native languages during word recognition, capitalising on the role of *cognates*.

Many languages share, to some extent, similarities at the lexical level. This is frequently due to their typological closeness and/or socio-historical events involving the speakers of these languages (e.g., migration, social contact). Cognates embody a particular case of such commonalities. Cognates are defined as cross-language synonyms which have similar forms whose form is similar (at the phonological level, orthographic level, signed level), commonly due to their shared etymological origin[1]. For example, Romance languages such as Spanish and Catalan share many cognates (Schepens et al., 2012), as in the case of *puerta* and *porta* (door in Spanish and Catalan, respectively). Cognates play a pivotal role in many models of bilingual lexical processing because they provide evidence that bilinguals access the lexicon in a language non-selective way. For instance, pictures are named faster by Spanish-Catalan bilinguals when their associated labels in both languages are cognates (e.g., *puerta-porta*), compared to when their labels are non-cognates (*mesa-taula*, [table]), which suggests that word representations of both languages are activated during word retrieval (Costa et al., 2000). This cognate advantage is not restricted to production, but also extends to word recognition (Midgley et al., 2011; Thierry & Wu, 2007), word learning (De Groot & Keijzer, 2000; Elias & Degani, 2022; Lotto & De Groot, 1998; Valente et al., 2018), and word translation (Christoffels et al., 2006).

1.1 Lexical and conceptual routes for translation

Early theoretical accounts of this 'Cognate effect' proposed that the impact of cognateness on lexical processing depends on bilinguals' proficiency in their second language (L2) (Potter et al., 1984; see Chen & Leung, 1989 for similar results). The reasoning behind this claim was the assumption that second-language learners start acquiring words in their L2 by associating them to their translation in their L1 (lexical route), instead of directly to their shared concept (conceptual route). This word-to-word connectivity should be sensitive to form-similarity (i.e., cognateness): the more similar the two word- forms are, the stronger the connection. As learners become more proficient, the connection between L2 representations and their meaning grows stronger, and L2 word processing becomes less reliant on word-word connections between L1 and L2 representations. Cognates subsequently exert less impact on lexical processing in proficient speakers (see Andras et al., 2022 for recent experimental evidence for this claim). The Revised Hierarchical Model (RHM, Kroll & Stewart, 1994) captured this hypothesis and predicted that translating words from L1 to L2 should take longer than translating from L2 to L1. The rationale behind this prediction was that L1-to-L2 translation

---

[1]Some form-similar cross-language synonyms are technically not cognates. For example, sun and sol (in Spanish) share phonological onset but their etymology points to different origins. The complimentary case is equally problematic when defining cognateness: cross-language synonyms that share etymological origin may not be necessarily be recognised as cognates by the average listener. Such is the case of cheese and queso which share Latin root but are phonologically dissimilar. For simplicity, we will use the term cognateness in this paper to include all cross-language synonyms that fulfill a specified threshold of form similarity, regardless of etymology. We do not expect that etymology plays a direct role on language perception if it is not via form-similarity, since it is not necessary for participants in psycholinguistic experiments to be aware of the etymology of the words they encounter in the tasks to be subject to the effect of form-similarity.

78 relies more strongly on word-to-word links between L1 and L2 representations, where L2-to-L1 translation
79 relies more strongly on the mediation between the concept and the two word- forms. L1-to-L2 translation
80 should therefore be more sensitive to the form similarity between the L1 and the L2 representations: cognate
81 words should be retrieved faster than non-cognate words during L1-to-L2 translation.

82 To test these predictions, Groot et al. (1994) and Groot (1992) asked Dutch natives with high (but non-
83 native) English proficiency to translate Dutch words to English (L1-to-L2) or English words to Dutch (L2-
84 to-L1). Although participants' performance in both conditions was roughly equivalent for cognates, non-
85 cognates were translated more slowly from Dutch to English than from English to Dutch, suggesting that
86 L1-to-L2 translation is more sensitive to form-similarity than L2-to-L1 translation. Conversely, participants'
87 performance was more sensitive to semantic variables (e.g., concreteness) when translating English words
88 to Dutch, compared to when translating Dutch words to English, suggesting that L2-to-L1 translation was
89 more sensitive to the shared conceptual properties of the translation pair.

90 These findings support a soft version of the RHM model, in which both the conceptual and the lexical
91 translation routes are active during translation, but where L1-to-L2 translation relies more strongly on
92 word-to-word links between L1 and L2 representations than L2-to-L1 translation does, and therefore is
93 more sensitive to cognateness. Further evidence for the presence of this difference between L1-to-L2 and
94 L2-to-L1 translations on the degree of reliance on the lexical and or the conceptual routes between L1-to-L2
95 and L2-to-L1 translation was provided by Phillips et al. (2006), who used eletrophysiological recordings of
96 bilinguals translating words from L1 to L2 or *vice versa* after semantic or phonological priming. The authors
97 reported stronger N400 responses (associated with semantic priming) in L1-to-L2 translation than in L2-
98 to-L12 translation, and stronger phonological mismatch negativity responses (associated with phonological
99 priming) during L2-to-L1 translation, compared to L1-to-L2 translation.

1.2 Competition from lexical neighbours

101 Since the RHM model was proposed, later studies on monolingual populations have focused how on the role
102 of the network of connections that words establish networks of connections with each other at the form-
103 level (phonological or orthographic) or at the conceptual level. For instance, the Neighbourhood Activation
104 Model (NAM, Luce & Pisoni, 1998) highlighted how lexical selection is affected by the number of phonological
105 neighbours that the selected word has. Words surrounded by a larger number of phonological neighbours
106 (words that only differ in one phoneme from the target word) are responded to more slowly and less accurately
107 than words from sparser phonological neighbourhoods (Goldinger et al., 1989; Luce et al., 1990), especially
108 when these neighbours' lexical frequency is higher (Luce & Pisoni, 1998).

109 The existence of phonological links between words is supported with experimental evidence from the visual
110 world paradigm and phonological priming. Meyer & Damian (2007) showed that target pictures were named
111 faster when presented together with a phonologically-related distractor, suggesting that the activation of
112 shared phonological information allowed the target to be activated and produced faster. there are also
113 findings suggesting that facilitatory or inhibitory effects on target recognition depend on the degree of
114 similarity the competitor shares with the target (Dufour & Peereman, 2003; Hamburger & Slowiaczek, 1996).
115 When there is only mismatch of one phoneme between prime and target, inhibitory effects are exerted on
116 target recognition (Dufour & Peereman, 2003). This inhibition was proposed to be due to strong lexical
117 activation of the prime which results in lexical interference, as opposed to simple phonological activation
118 from less similar primes that triggers facilitation.

119 These findings of phonological links between words in the monolingual lexicon pose an important question in
120 the field of bilingualism research: do word representations in one language form part of the phonological or
121 orthographic neighbours of the other language? Weber & Cutler (2004) observed that bilinguals displayed
122 lexical competition effects from words in their native language when performing a word-referent matching
123 task in their L2. This highlights that lexical competition for bilinguals during language comprehension may
124 not only come from the tested language but also the native language, despite the native language not being
125 actively required for the task. Other studies have similarly found interference when spoken target words in
126 one language had phonological overlap with a word in the participant's other language, even when the other
127 language was not required by the task (Desroches et al., 2022; Spivey & Marian, 1999). Such patterns have
128 even been found with young children less than 4 years old (Von Holzen & Mani, 2012).

129 The Bilingual Interactive Activation (BIA/BIA+) (Dijkstra & Van Heuven, 2002; Van Heuven et al., 1998)
130 addressed this issue and suggested that lexical representations from both languages establish both excitatory
131 and inhibitory connections with each other, resulting in an integrated lexicon for both languages. This
132 is, in line with connectionist approaches to lexical processing (McClelland & Rumelhart, 1981). Evidence
133 supporting such an integrated lexicon is provided by the fact that bilinguals' performance in word recognition

134 tasks is sensitive to the orthographic neighbourhood density of the presented word's translation. When
135 bilingual participants are presented with words in one language in a lexical decision task, translating them
136 to the other language took longer when they were part of large neighbourhoods (Van Heuven et al., 1998).
137 This suggests that presented words activated orthographic neighbourhoods in the other language, which
138 competed for selection with the target word during recognition.

1.3   The present study

140 In the present study, we tested listeners' reliance on phonological similarity when translating words from
141 an unfamiliar language to their native language, in order to investigate the plausibility of word-to-word
142 connections during L1-to-L2 translation. We tested English native speakers and Spanish native speakers in
143 a translation elicitation task. This paradigm has been used by bilingualism researchers to identify easily-
144 recognisable cognates for experimental studies. Using this paradigm, we aim to delve deeper into the factors
145 affecting foreign word recognition by monolingual participants with no prior exposure to the language.

146 Participants were presented with spoken words in an unfamiliar (non-native language) and were asked to
147 translate them to their native language. We will henceforth refer to the auditory-presented words heard
148 by participants on each trial as presented words, and the correct translation for the presented words as
149 target translations. The motivation behind testing participants in an unfamiliar language was that these
150 participants can be considered a particular (extreme) case of unbalanced bilingualism, in which they do not
151 know any words in L2, and therefore word-to-concept connections are, in principle, absent. As a consequence,
152 these participants should only be able use word-to-word connections to succeed in the translation of words
153 in an unfamiliar language. We predicted that, if the information provided by the phonological form of the
154 unfamiliar presented words is sufficient for translation, participants' performance should increase when the
155 translation pairs are phonologically more similar.

156 Following Luce & Pisoni (1998)'s NAM model, we further hypothesised that if participants are able to
157 correctly translate words from an unfamiliar language based on their phonological similarity to the target
158 word, similar sounding but incorrect translations (e.g., false friends) should be activated too. If this is the
159 case, non-native words that have more phonological neighbours should be less likely to be correctly translated,
160 especially if such neighbours have higher lexical frequency and are phonologically closer to the presented word
161 than the target translation. For instance, one could expect English participants to incorrectly translate the
162 Spanish word *botón* as bottom instead of as its correct translation *button* (because the lexical frequency
163 of *bottom* is higher than that of *button* and also phonological closer). This combined constraint led us to
164 calculate a specific form of phonological neighbours where a neighbour is counted only if it is higher frequency
165 and has stronger phonological similarity than the target translation as judged by Levenshtein distance. If
166 the phonological neighbourhood density of the target translation affects participants' performance negatively,
167 this would suggest that non-native words trigger competition between words in the native language during
168 translation.

169 In Experiment 1, we collected data from two groups of British English natives. One group was presented
170 with Catalan words, the other with Spanish words. We examined the extent to which participants were
171 able to use the phonological similarity between the presented (Catalan or Spanish) word and its translation
172 (i.e., cognateness) to provide accurate responses. In Experiment 2, we tested a group of (European) Spanish
173 natives in the same task, who were presented with Catalan words. Catalan and Spanish are both Romance
174 languages whose close typological distance is reflected in the fact that they share many cognates, where
175 English is a Germanic language that shares considerably fewer cognates with Catalan and Spanish. By testing
176 participants translating words from typologically close and distant languages, we expected to widen the range
177 of the phonological similarity scores of the translation pairs involved in the experimental task, therefore
178 allowing us to explore potential cross-language differences in participant's performance. Participants in
179 Experiments 1 and 2 were surprisingly good at translation some words, which had low phonological similarity
180 with their correct translation. In Experiment 3, we collected additional data from a new group of British
181 English natives, who were presented with Catalan and Spanish words. In contrast with the procedure in
182 Experiment 1, participants in Experiment 3 where asked, after providing a response in each trial, to report
183 whether they had previous knowledge of the meaning of the presented word (e.g., some Spanish words are
184 common knowledge, even for monolingual speakers of English). Finally, we present analyses on the joint
185 data sets of Experiments 1 to 3.

2 Experiment 1

2.1 Methods

All materials, data, and code used in this study are hosted in the Open Science Framework https://osf.io/9fjxm/ and a GitHub repository https://github.com/gongcastro/translation-elicitation.git, along with additional notes.

2.1.1 Participants

Participants were recruited via Prolific (5£ compensation) and SONA (compensation in academic credits), and gave informed consent before providing any data and the study was conducted in accordance with ethical standards of the Declaration of Helsinki and the protocol was by the University of Oxford Medical Sciences Inter-Divisional Research Ethics Committee (IDREC) (R60939/RE007). Participants were asked to complete the experiment using a laptop in a quiet place with good internet connection. Data collection took place from June 04th, 2020 to June 25th, 2020. We collected data from 71 British English native participants living in United Kingdom (*Mean* = 21.76 years, *SD* = 2.15, Range = 18-26, 46, 25 female).

Table 1

| | | Age | | |
	N[1]	Mean ± SD	Range	L2
Experiment 1				
spa–ENG	35 (8)	21.80 ± 2.08	18–26	Russian (1)
cat–ENG	36 (4)	21.72 ± 2.25	18–25	French (1), German (1), Italian (1), Punjabi (1), Several (1)
Experiment 2				
cat–SPA	33 (12)	21.85 ± 3.00	18–33	French (9), German (1), Italian (2)
Experiment 3				
spa–ENG	32 (1)	21.72 ± 2.59	18–26	German (2), Japanese (1)
cat–ENG	32 (0)	22.31 ± 2.39	18–26	Cantonese (1), Irish (1)

[1]Number of included participants (number of excluded participants.)

2.1.2 Stimuli

We arranged two lists of input words to be presented to participants in the auditory modality: one in Catalan and one in Spanish. Words in the Catalan list were 5.02 phonemes long on average (*SD* = 1.49, Range = 2-8), and the orthographic forms of their English translations (which participants had to type) were 5.12 characters long on average (*SD* = 1.56, Range = 3-9). Words in the Spanish list were 5.52 phonemes long on average (*SD* = 1.47, Range = 3-9), and the orthographic form of their English translations (which participants had to type), were 5.29 characters long on average (*SD* = 1.77, Range = 3-12).

Participants listened to one audio file in each trial, each containing a single word presented in isolation. The audio files were the same ones used in child experiments conducted in the Laboratori de Recerca en Infància of Universitat Pompeu Fabra (Barcelona, Spain). These audio files were recorded by a proficient Catalan-Spanish female bilingual from the Metropolitan Area of Barcelona in a child-directed manner. Catalan and Spanish words were recorded at 44,100 Hz in separate files in the same session, and then de-noised using Audacity and normalised at peak intensity using Praat (Broersma & Weenink, 2021). The average duration of the Catalan audio files was 1.24 (*SD* = 0.19, Range = 0.8-1.58). The average duration of the Catalan audio files was 1.16 (*SD* = 0.15, Range = 0.78-1.53).

For each translation pair, we defined three predictors of interest: the lexical frequency of the correct translation (*Frequency*), the phonological similarity between the presented (Catalan or Spanish) word and their correct English translation (*Cognateness*), and the presented word's number of cross-language phonological neighbours (*CLPN*).

We included *Frequency* as a nuance predictor, under the hypothesis that—keeping other predictors constant— participants would translate higher-frequency words more accurately and faster than lower-frequency words. Lexical frequencies of correct translations were extracted from SUBTLEX-UK (Van Heuven et al., 2014), and transformed to Zipf scores. Words in the stimuli list without a lexical frequency value were excluded from data analysis (2 in the Catalan list, 5 in the Spanish list).

224 We calculated *Cognatenesss*, our main predictor of interest, by computing the Levenshtein similarity between
225 the X-SAMPA transcriptions of each pair of translations using the `stringdist` R package (van der Loo, 2014).
226 The Levenshtein distance computes the edit distance between two character strings—in this case, phoneme
227 transcriptions—by counting the number of additions, deletions, and substitutions necessary to make both
228 strings identical (Levenshtein et al., 1966). We divided this edit distance by the number of characters included
229 in the longest X-SAMPA transcription of the translation pair. This results in a proportion score, in which
230 values closer to zero correspond to lower Levenshtein distances between phonological transcriptions (i.e.,
231 higher similarity), and values closer to 1 correspond to higher Levenshtein distances (i.e., lower similarity).
232 This transformation also help account for the fact that the Levenshtein distance tends to increase with the
233 length of the transcriptions. For interpretability, we subtracted this proportion from one, so that values
234 closer to one correspond to higher similarity between phonological transcriptions, and values closer to zero
235 correspond to lower similarity between phonological transcriptions. For example, the *table* (te bəl)-*mesa*
236 (mesa) translation pair had a 17% similarity, while the *train* (`trEIn`)-*tren* (t En) translation pair had a 60%
237 similarity. Figure 1 summarises the lexical frequency, phonological neighbourhood density and phonological
238 overlap of the words included in the Catalan and the Spanish lists.

239 For each Catalan and Spanish word, we calculated the number of *CLPN* by counting the number of English
240 words with same or higher lexical frequency, and whose phonological transcription (in X-SAMPA format)
241 different in up to one phoneme from that of the presented Catalan or Spanish word. Lexical frequencies
242 and phonological transcriptions were extracted from the multilingual database CLEARPOND (Marian et
243 al., 2012)[2].

Table 2Stimuli properties.

	Frequency		Cognateness		CLPN	
	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range
spa-ENG	5.88 ± 0.57	4.43–7.24	0.13 ± 0.18	0.00–0.75	0.19 ± 0.79	0–5
cat-ENG	5.89 ± 0.59	4.43–7.27	0.16 ± 0.18	0.00–0.67	0.76 ± 2.53	0–15
cat-SPA	5.79 ± 0.63	4.48–7.70	0.38 ± 0.26	0.00–1.00	0.87 ± 2.08	0–12
Mean	5.85	—	0.225	—	0.606	—

2.1.3 Procedure

245 The experiment was implemented online using Psychopy/Pavlovia (Peirce et al., 2019). Participants accessed
246 the study from a link provided by Prolific or SONA and completed the experiment from an internet browser
247 (Chrome or Mozilla). After giving their consent for participating, participants answered a series of questions
248 about their demographic status, their language background, and the set up they were using for completing
249 the study. Then, participants completed the experimental task. Participants were informed that they would
250 listen to a series of pre-recorded words in either Catalan or Spanish (English participants) or Catalan (Spanish
251 participants). They were instructed to listen to each word, guess its meaning in English (English participants)
252 or Spanish (Spanish participants), and type their answer as soon as possible. English participants were
253 randomly assigned to the Catalan or Spanish lists. Participants in the Catalan list were presented with 83
254 trials, and participants in the Spanish list were presented with 99 trials.

255 Each trial started with a yellow fixation point presented during one second on the centre of the screen over
256 a black background. After one second, the audio started playing while the dot remained being displayed
257 until the audio ended. Upon the offset of the fixation point and audio, participants were prompted to
258 write their answer by a `>` symbol. Typed letters were displayed in the screen in real time to provide visual
259 feed-back to participants. Participants were allowed to correct their answer. Then, participants pressed the
260 `RETURN/ENTER` key to confirm their answer and start and new trial.

---

[2][Phonological transcriptions in CLEARPOND were generated from eSPEAK, http://espeak.sourceforge.net/]
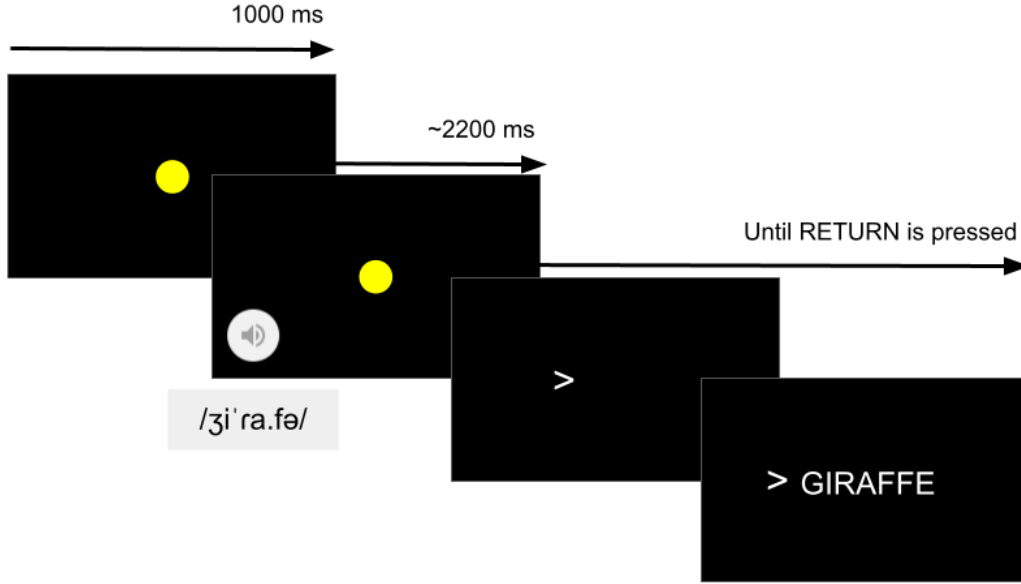
Figure 1Schematic representation of a trial in the experimental task

2.1.4 Data analysis

2.1.4.1 Data processing After data collection, participants answers were manually coded into the following categories: *Correct, Typo, Wrong, False friend, Other*. Responses were coded as *Correct* if the provided string of characters was identical to the orthographic form of the correct translation. Trials were coded as *Typo* if the participant provided a string of characters only one edit distance (addition, deletion, or substitution) apart from the orthographic form of the correct translation (e.g., "pengiun" instead of "penguin"), as long as the response did not correspond to a distinct English word. Responses were coded as *False friend* if the participant provided a phonologically similar incorrect translation. Responses were coded as *Other* (see Data analysis section for more details). Both *Correct* and *Typo* responses were considered as correct, while *Wrong* and *False friend* responses were considered as incorrect. *Other* responses were excluded from data analysis.

Trials in which participants took longer than 10 seconds to respond were also excluded. Participants contributed a total of 9152 valid trials (5694 in Catalan, 3458 in Spanish). The task took approximately 15 minutes to be completed.

2.1.4.2 Modelling approach and statistical inference We modelled the probability of participants guessing the correct translation of each presented word using a generalised multilevel Bayesian regression model with a Bernoulli logit link distribution. We included as fixed effects the intercept, the main effects of *Frequency*, *Cognateness*, and *CLPN*, and the two-way interaction between *Cognateness* and *CLPN*. We also included participant-level random intercepts and slopes for the the main effects and the interaction. Eq. 1 shows a formal description of the model.

Likelihood:

$$y_i \sim \text{Bernoulli}(p_i)$$

Linear model:

$$\text{logit}(p_i) = \beta_{0_j} + \beta_{1_j}\text{Frequency} + \beta_{2_j}\text{CLPN} + \beta_{3_j}\text{Cognateness} +$$
$$\beta_{4_j}(\text{CLPN}_i \times \text{Cognateness})$$

$$(1)$$

$$\beta_{0-4_j} \sim \mathcal{N}(\mu_{\beta_j}, \sigma_{\beta_j}), \text{ for participant } j$$

Prior:

$$\mu_{\beta_p} \sim \mathcal{N}(0, 0.1)$$
$$\sigma_{\beta_p} \sim \text{HalfCauchy}(0, 0.1)$$
$$\rho_p \sim LKJ(8)$$

Figure 2Eq. 1

To test the practical relevance of each predictor we followed (Kruschke & Liddell, 2018). We first specified a region of practical equivalence (*ROPE*) around zero ([-0.1, +0.1], in the logit scale). This area indicates the values of the regression coefficients that we considered as equivalent to zero. We then computed the 95% posterior credible intervals (CrI) of the regression coefficients of interest. Finally, we calculated the proportion of the 95% CrI that fell inside the ROPE. This proportion indicates the probability that the true value of the coefficient is equivalent to zero.

All analyses were performed in R environment (RCore, 2019). We used the tidyverse family of R packages (Wickham et al., 2019) to process data and to generate figures. We used the `brms` R package (Bürkner, 2017) using the `cmdstanr` back-end to the Stan probabilistic language (Carpenter et al., 2017) to estimate and compare the models (see Appendix 1 for mode details on the models).

2.2 Results

We collected data for a total of 6,446 trials completed by 72 unique participants. Of those trials, 2,988 were provided by 36 unique participants who listened to Catalan words, and 3,458 trials were provided by 36 unique participants who listened to Spanish words. We excluded trials in which participants did not enter any text ($n = 72$), in which a response in a language other than English was provided (e.g., `agua`, $n = 51$), in which participants did not provide a whole word (e.g., `f`, $n = 5$), and in which participants added comments to the experimenter (e.g., `unsure`, ($n = 13$). In addition, we excluded data from participants that self-rated their oral and/or written skills in Catalan and Spanish, or any other second language as four or higher in a five-point scale ($n = 2$), were diagnosed with a language ($n = 1$), or did not contribute more than 80% of valid trials ($n = 9$).

After applying trial-level and participant-level inclusion criteria, the resulting dataset included 5,204 trials provided by 54 unique participants. Of those trials, 2,602 were provided by 27 unique participants who listened to Catalan words, and 2,604 trials were provided by 32 unique participants who listened to Spanish words. Responses given by English participants to Catalan presented words were 5.35 characters long on average (*Median* = 5, *SD* = 1.79, Range = 1-14), while their translations to Spanish responses were 5.57 characters long on average (*Median* = 5, *SD* = 1.97, Range = 2-21).

Table 3

	Correct responses				Incorrect responses			
	Correct	(%)	Typo	(%)	Wrong	(%)	False friend	(%)
Experiment 1								
cat-ENG	429	(16.47)	11	(0.42)	2,082	(79.95)	82	(3.15)
spa-ENG	374	(14.37)	2	(0.08)	2,117	(81.36)	109	(4.19)
Sum	803	—	13	—	4,199	—	191	—
Mean	—	15.42	—	0.25	—	80.66	—	3.67
Experiment 2								
cat-SPA	780	(46.93)	20	(1.20)	736	(44.28)	126	(7.58)
Sum	780	—	20	—	736	—	126	—
Mean	—	46.93	—	1.20	—	44.28	—	7.58
Experiment 3								
cat-ENG	477	(18.01)	7	(0.26)	1,986	(75.00)	178	(6.72)
spa-ENG	590	(19.53)	6	(0.20)	2,294	(75.94)	131	(4.34)
Sum	1,067	—	13	—	4,280	—	309	—
Mean	—	18.77	—	0.23	—	75.47	—	5.53

Table 3 shows a summary of participants' accuracy across Experiments 1, 2, and 3. MCMC chains in the model showed strong evidence of convergence ($\hat{R} < 1.01$) (see Appendix 2 for more detailed model diagnostics). Participants translating Catalan words and participants translating Spanish words performed equivalently, as indicated by the regression coefficient of *Group* (β = -0.199, 95% CrI = [-0.512, 0.129], p(ROPE) = 0.238). Overall, participants responded less accurately to words with more CLPNs than to words with fewer CLPNs, regardless of the amount of phonological similarity between the presented word and its translation. This is indicated by the size of the regression coefficient of the two-way interaction between *Cognateness* and *CLPN* (β = -0.653, 95% CrI = [-0.973, -0.313], p(ROPE) = 0). As anticipated, participants' performance benefited from an increase in *Cognateness* (β = 7.133, 95% CrI = [6.606, 7.735], p(ROPE) = 0), while the number of *CLND* had the opposite effect (β = 0.009, 95% CrI = [-0.11, 0.111], p(ROPE) = 0.925). Figure 3 illustrates the posterior of the average predictions of the model for words with different values of *Cognateness* and *CLPN*. Figure 7 shows a graphic summary of the posterior distribution of the regression coefficients of interest.

Table 4

		Accuracy (%)				Valid trials			
	N	Mean	SD	SE	Range	Mean	N trials	SD	Range
Experiment 1									
spa-ENG	27	15.86	5.20	3.05	8.82–28.71	96.37	2,602	2.88	87–98
cat-ENG	32	18.48	4.89	3.27	10.84–32.56	81.38	2,604	3.17	71–83
Sum	59	—	—	—	—	—	5,206	—	158
Mean	—	17.17	5.05	3.16	—	8,887.27	—	302.71	—
Experiment 2									
cat-SPA	21	48.30	5.29	10.54	38.27–58.97	79.14	1,662	3.14	72–82
Sum	21	—	—	—	—	—	1,662	—	72
Mean	—	48.30	5.29	10.54	—	7,914.29	—	313.51	—
Experiment 3									
spa-ENG	31	20.92	8.29	3.76	5.88–44.12	97.45	3,021	1.80	88–98
cat-ENG	32	19.74	4.94	3.49	10.34–27.91	82.75	2,648	0.44	82–83
Sum	63	—	—	—	—	—	5,669	—	170

9

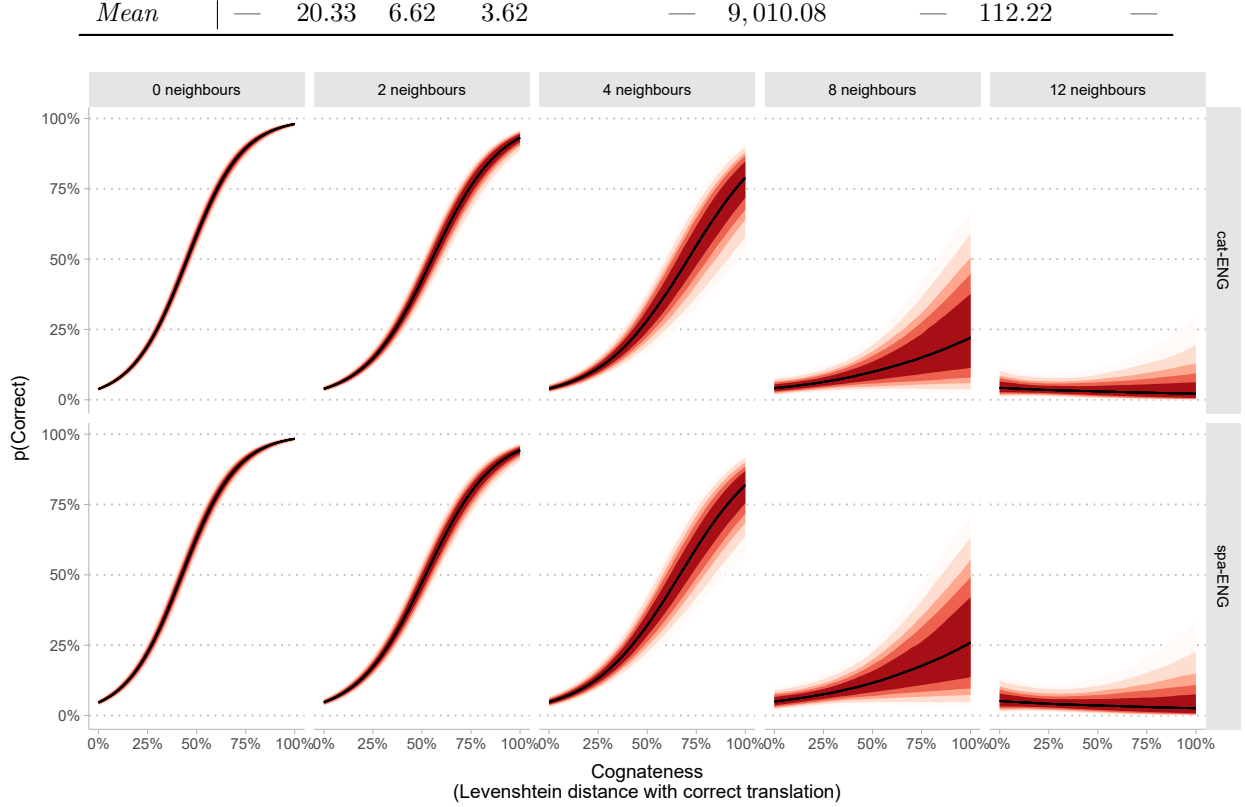| Mean | | — | 20.33 | 6.62 | 3.62 | — | 9,010.08 | — | 112.22 | — |

Figure 3 Posterior model-predicted mean accuracy in Experiment 1. Predictions were generated from 4,000 posterior samples, extracted for different values of CLPN (0, 2, 4, 8, 12) and Cognateness (1-100). Predictions are plotted separately for English participants translating Catalan words, and for English participants translating Spanish words. Lines indicate mean predictions, and intervals indicate 95%, 89%, 78%, 67%, and 50% credible intervals (CrI).

2.3 Discussion

In the present experiment, we investigated the extent to which cognateness (phonological similarity between translation equivalents) is sufficient for successful word translation, in the absence of semantic knowledge. We tested two groups of monolingual British English natives in a translation task that involved words in Catalan or Spanish, two languages participants reported having no prior familiarity with. Participants benefited strongly from cognateness, when the correct translation of the presented words in Catalan or Spanish had few English phonological neighbours with higher lexical frequency. This suggests that, in the absence of distractors, even naïve participants can efficiently use phonological similarity to succeed in a translation task. Our results suggest that word-to-word connections at the phonological level might play a strong role during L2 speech comprehension, specially in low-proficiency listeners.

3 Experiment 2

Results in Experiment 1 suggest that English natives were able to exploit the phonological similarity between unfamiliar words in Catalan and Spanish to provide accurate translations to English. English, a Germanic language, is relatively distant from Catalan and Spanish, two Romance languages. English shares fewer cognates with Catalan and Spanish than it does with typologically closer languages, like Dutch and German. In Experiment 2, we investigated whether listeners of an unfamiliar but typologically closer language benefit more strongly from phonological similarity when performing the same task as in Experiment 1. To this aim, we presented Spanish participants, who reported little-to-no prior familiarity with Catalan, with Catalan words.

3.1 Methods

Participants in Spain were contacted via announcements in Faculties, and were compensated 5€ or an Amazon voucher for the same value. Participants gave informed consent before providing any data and the

study was conducted in accordance with ethical standards of the Declaration of Helsinki and the protocol was approved by the Drug Research Ethical Committee (CEIm) of the IMIM Parc de Salut Mar (2020/9080/I). Data collection took place from June 08th, 2020 to June 28th, 2020. We collected data from 33 Spanish native participants living in Spain (*Mean* = 21.85 years, *SD* = 3, Range = 18-33, 28, 5 female). Stimuli were the same list of Catalan stimuli as in Experiment 1. Procedure and data analysis were identical as in Experiment 1.

3.2 Results

We collected data for a total of 5,412 trials completed by 33 unique participants. We excluded trials in which participants did not enter any text ($n = 44$), in which a response in a language other than English was provided (e.g., `agua`, $n = 51$), in which participants did not provide a whole word (e.g., `f`, $n = 7$), and in which participants added comments to the experimenter (e.g., `unsure`, ($n = 1$). In addition, we excluded data from participants that self-rated their oral and/or written skills in Catalan and Spanish, or any other second language as four or higher in a five-point scale ($n = 22$), were diagnosed with a language ($n = 1$), or did not contribute more than 80% of valid trials ($n = 9$). After applying trial-level and participant-level inclusion criteria, the resulting dataset included 1,662 trials provided by 42 unique participants. Of those trials, 1,662 were provided by 21 unique participants who listened to Catalan words, and 1,662 trials were provided by 21 unique participants who listened to Spanish words. Responses given by participants were 5.6 characters long on average (*Median* = 5, *SD* = 1.6, Range = 2-12).

MCMC chains in the model showed strong evidence of convergence ($\hat{R} < 1.01$) (see Appendix 2 for more detailed model diagnostics). Overall, participants responded less accurately to words with more CLPNs than to words with fewer CLPNs, regardless of the amount of phonological similarity between the presented word and its translation. This is indicated by the size of the regression coefficient of the two-way interaction between *Cognateness* and *CLPN* (β = -0.409, 95% CrI = [-0.895, 0.111], p(ROPE) = 0.081). As anticipated, participants' performance benefited from an increase in *Cognateness* (β = 9.274, 95% CrI = [8.473, 10.269], p(ROPE) = 0), while the number of *CLND* had the opposite effect (β = -0.069, 95% CrI = [-0.353, 0.192], p(ROPE) = 0.492). Figure 3 illustrates the posterior of the average predictions of the model for words with different values of *Cognateness* and *CLPN*.
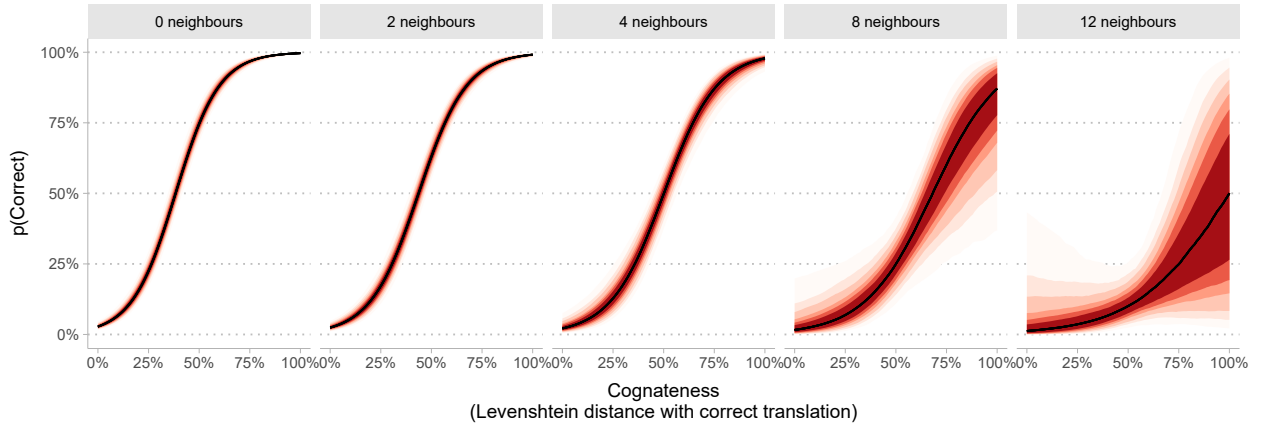

Figure 4Posterior model-predicted mean accuracy in Experiment 1. Predictions were generated from 4,000 posterior samples, extracted for different values of CLPN (0, 2, 4, 8, 12) and Cognateness (1-100). Predictions are plotted separately for English participants translating Catalan words, and for English participants translating Spanish words. Lines indicate mean predictions, and intervals indicate 95%, 89%, 78%, 67%, and 50% credible intervals (CrI).

3.3 Discussion

Experiment 2 was an extension of Experiment 1 to a population of monolinguals whose native language is typologically similar to the presented language. Particularly, we presented Catalan words to Spanish native speakers who were reportedly unfamiliar with Catalan. Our results indicate a similar pattern of results as those in Experiment 1: participants were able to provide correct translations of presented Catalan words, provided that the Catalan words shared some degree of phonological similarity with their Spanish translation, and that the number of phonological neighbours with higher lexical frequency was reduced. In contrasts with

375 the results in Experiment 1, the positive impact of phonological similarity on participants' performance in
376 Experiment 2 was more resilient to the interference of phonological neighbours. English natives in Experiment
377 1 provided significantly less accuracy responses when more than four phonological neighbors were present
378 (even when translating high-similarity words), compared to when only one or neighbour were present. Spanish
379 participants in Experiment 2 benefited from phonological similarity, even when eight neighbours were present.
380 Spanish participants' performance declined after 8 neighbours, and was evident at 12 neighbours. Overall,
381 this suggests that participants in Experiment 2, who were natives of a typologically similar language (Spanish)
382 to the presented language (Catalan) benefited more strongly from phonological similarity than participants
383 in Experiment 1, who were natives of typologically less similar language (English) to the presented language
384 (Catalan, Spanish).

385 Participants from both Experiment 1 and 2 benefited strongly from phonological similarity to correctly
386 translate words from a non-native, reportedly unfamiliar language. This pattern of results holds for most
387 of the presented stimuli, but some low-similarity Catalan and Spanish words were responded to surprisingly
388 accurately by English listeners. Given that participants were reportedly unfamiliar with both languages, it
389 was expected that participants would be very unlikely to provide correct translations for words sharing little
390 to no phonological similarity to their correct translation. Table 5 a list of Catalan and Spanish words to
391 which participants provided responses with ≥ 10 average accuracy.

Table 5List of items with unexpectedly high accuracy: the Levenshtein similarityscore betwen th presented word (in Catalan or Spanish) and their correct Enlgish or Spanish translation is zero, but participants, who are reportedly unfamiliar with the presented language, were on average >10% likely to guess the correct translation.

	Accuracy (%)	SE
Experiment 1 - cat-ENG		
cavall /kə a / - horse /h s/	17.14	6.37
llibre / i ə/ - book /b k/	17.14	6.37
camisa /kəmizə/ - shirt / t/	16.67	6.21
poma /poma/ - apple /æpl/	16.67	6.21
cama /kamə/ - leg /l g/	11.11	5.24
Experiment 1 - spa-ENG		
pantalón /paɲtalon/ - trousers /tra zəz/	77.42	7.51
naranja /na aŋxa/ - orange / r n /	41.94	8.86
leche /le e/ - milk /m lk/	35.48	8.59
toro /to o/ - bull /b l/	33.33	8.61
libro /li o/ - book /b k/	30.00	8.37
cebra / eb a/ - zebra /zibrə/	29.03	8.15
pan /pan/ - bread /br d/	29.03	8.15
pollo /po o/ - chicken / k n/	26.67	8.07
jirafa /xi afa/ - giraffe / r f/	20.69	7.52
perro /pero/ - dog /d /	16.13	6.61
pluma /pluma/ - feather /f ðə/	16.13	6.61
puerta /pwerta/ - door /d /	16.13	6.61
pie /pje/ - foot /f t/	12.90	6.02
caballo /ka a o/ - horse /h s/	10.34	5.66
bocadillo /bokadi o/ - sandwich /sænw /	10.00	5.48
globo / lo o/ - balloon /bəlun/	10.00	5.48
Experiment 2 - cat-SPA		
fulla /fu ə/ - hoja /oxa/	30.43	9.59
ull /u / - ojo /oxo/	21.74	8.60
got /g t/ - vaso /baso/	20.00	8.00
entrepà /ent əpa/ - bocadillo /bokadi o/	13.04	7.02
mirall /mi a / - espejo /espexo/	12.50	6.75

392 It is likely that participants had prior knowledge of these words despite having reported little to no famil-
393 iarity with the presented language (Catalan or Spanish). One possibility is that participants had previously

encountered these words embedded in English linguistic input. Spanish words percolate English speech with relative frequency, via different sources such as popular culture, songs, TV programs, etc. In addition, words from languages other than Spanish, but with high similarity to the Spanish words (e.g., cognates from Italian or French) might appear in English speech as well. Such prior knowledge might not be specific to the low-similarity words highlighted before. Participants may also have had prior knowledge about higher-similarity words, which could have contributed to participants responding to such words more accurately than without such prior knowledge. In the case of higher-similarity words, it is more difficult to disentangle the extent to which participants' accuracy is a function of pure phonological similarity, or prior knowledge they had about the meaning of Spanish words. To investigate this issue, we run Experiment 3.

[ADD HERE LEXICAL FREQUENCY OF CATALAN AND SPANISH WORDS IN ENGLISH, AND OF CATALAN WORDS IN SPANISH]

4 Experiment 3

Experiment 3 is a replication of Experiment 1, in which we collected additional data about participants' prior familiarity with the presented Catalan and Spanish words, in addition to the same translation task presented to participants in Experiment 1.

4.1 Methods

Data collection took place from October 22th, 2022 to October 23th, 2022. We collected data from 64 British English native participants living in United Kingdom ($Mean = 22.02$ years, $SD = 2.49$, Range = 18-26, 36, 28 female).

Participants were recruited via Prolific (5£ compensation) and SONA (compensation in academic credits), and gave informed consent before providing any data and the study was conducted in accordance with ethical standards of the Declaration of Helsinki and the protocol was by the University of Oxford Medical Sciences Inter-Divisional Research Ethics Committee (IDREC) (R60939/RE007). Participants were asked to complete the experiment using a laptop in a quiet place with good internet connection. Stimuli were the same list of Catalan stimuli as in Experiment 1.

The experiment was implemented online using Qualtrics (Qualtrics, Provo, UT). This platform was chosen to allow easier presentation of survey questions aimed to probe prior understanding of the presented words and participants' confidence ratings of their answers. With the exception of these additional questions, we attempted to replicate the procedure of Experiment 1 as closely as possible. The Spanish and Catalan audio stimuli used were identical the materials in Experiment 1. Participants were randomly assigned to the Catalan or Spanish lists. The Catalan list had 83 trials and the Spanish list had 99 trials. Participants first completed the consent form followed by the questionnaire about demographic status, language background and set up. They then proceeded to the experimental task.

On each trial, participants listened to the audio stimulus by clicking on the play button. For comparability to the PsychoPy version, participants were only allowed to play the audio one time. Participants were explicitly told that they would be only allowed to listen once. The play button vanished after one playthrough. Participants then had to answer three questions based on the audio they had heard on that trial. These questions were presented on the same page, directly below the audio player. They were first asked whether or not they knew the presented word (multiple choice—yes/no). Regardless of their answer on the first question, participants were asked what they thought the translation of the word was in English (or their best guess), and instructed to type their answer in the provided text box. Finally, they were asked to rate how confident they were in their answer on a scale of 0 to 7, where 7 was "very confident" and 0 was "not confident". There was no time limit on the response phase. All questions had to be answered to proceed to the next trial.

Participants first completed 5 practice trials with English words as the audio stimulus (ambulance, cucumber, elephant, pear, turtle). The words were recorded by a female native speaker of English. These trials acted as attention checks, as participants should always answer "yes" to the first question on prior word knowledge and be able to accurately transcribe the word they heard. Following the practice phase, participants completed the test phase where they heard either Spanish words or Catalan words.

4.2 Results

We collected data for a total of 6,016 trials completed by 64 unique participants. Of those trials, 2,752 were provided by 32 unique participants who listened to Catalan words, and 3,264 trials were provided by 32 unique participants who listened to Spanish words. We excluded trials in which participants did not enter

any text ($n =$), in which a response in a language other than English was provided (e.g., agua, $n =$), in which participants did not provide a whole word (e.g., f, $n =$), and in which participants added comments to the experimenter (e.g., unsure, ($n =$). In addition, we excluded data from participants that self-rated their oral and/or written skills in Catalan and Spanish, or any other second language as four or higher in a five-point scale ($n = 2$), were diagnosed with a language ($n = 0$), or did not contribute more than 80% of valid trials ($n = 1$).

After applying trial-level and participant-level inclusion criteria, the resulting dataset included 6,290 trials provided by 62 unique participants. Of those trials, 3,145 were provided by 31 unique participants who listened to Catalan words, and 2,743 trials were provided by 32 unique participants who listened to Spanish words.

INSERT PRIOR KNOWLEDGE FILTER

Responses given by English participants to Catalan presented words were 5.52 characters long on average (*Median* = 5, *SD* = 1.75, Range = 1-17), while their translations to Spanish responses were 5.41 characters long on average (*Median* = 5, *SD* = 1.77, Range = 1-20).

MCMC chains in the model showed strong evidence of convergence ($\hat{R} < 1.01$) (see Appendix 2 for more detailed model diagnostics). Overall, participants reported prior knowledge more often for that Spanish words with unexpectedly high accuracy (see Discussion in Experiment 2) than for words with expected accuracy (see Figure 5). Participants reported prior knowledge of Catalan words with unexpected accuracy as often as those with expected accuracy. This suggests that participants in Experiment 1 may have relied, to some extent, on their prior knowledge about form-meaning mappings to correctly translate some Spanish words. To isolate such an effect of prior Spanish knowledge, we run the same analysis as in Experiment 1 on the newly collected translations from Experiment 3, now excluding responses to words in which participants reported prior knowledge.

Participants translating Catalan words and participants translating Spanish words, as indicated by the regression coefficient of *Group* (β = -0.049, 95% CrI = [-0.388, -0.32], p(ROPE) = 0.426). Overall, participants responded less accurately to words with more CLPNs than to words with fewer CLPNs, regardless of the amount of phonological similarity between the presented word and its translation. This is indicated by the size of the regression coefficient of the two-way interaction between *Cognateness* and *CLPN* (β = -0.665, 95% CrI = [-1.068, -0.32], p(ROPE) = 0.002). As anticipated, participants' performance benefited from an increase in *Cognateness* (β = 7.585, 95% CrI = [7.043, 8.24], p(ROPE) = 0), while the number of *CLND* had the opposite effect (β = -0.049, 95% CrI = [-0.203, 0.086], p(ROPE) = 0.753). Figure 3 illustrates the posterior of the average predictions of the model for words with different values of *Cognateness* and *CLPN*.

4.3 Discussion

5 Joint analyses

Across Experiments 1 and 3, we found strong evidence that participants efficiently exploited phonological similarity to provide accurate translations for words in an unfamiliar language, provided that few phonological neighbours of higher lexical frequency were present. Figure 7 summarizes the posterior distribution of the regression coefficients of the models in Experiments 1 to 3.

6 General discussion

7 References

Andras, F., Rivera, M., Bajo, T., Dussias, P. E., & Paolieri, D. (2022). Cognate facilitation effect during auditory comprehension of a second language: A visual world eye-tracking study. *International Journal of Bilingualism*, 13670069211033359.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, *109*(2), 775–794.

Broersma, P., & Weenink, D. (2021). *Praat: Doing phonetics by computer [computer program]* (Version 6.1.54). http://www.praat.org/

Bürkner, P.-C. (2017). Brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1), 1–28.
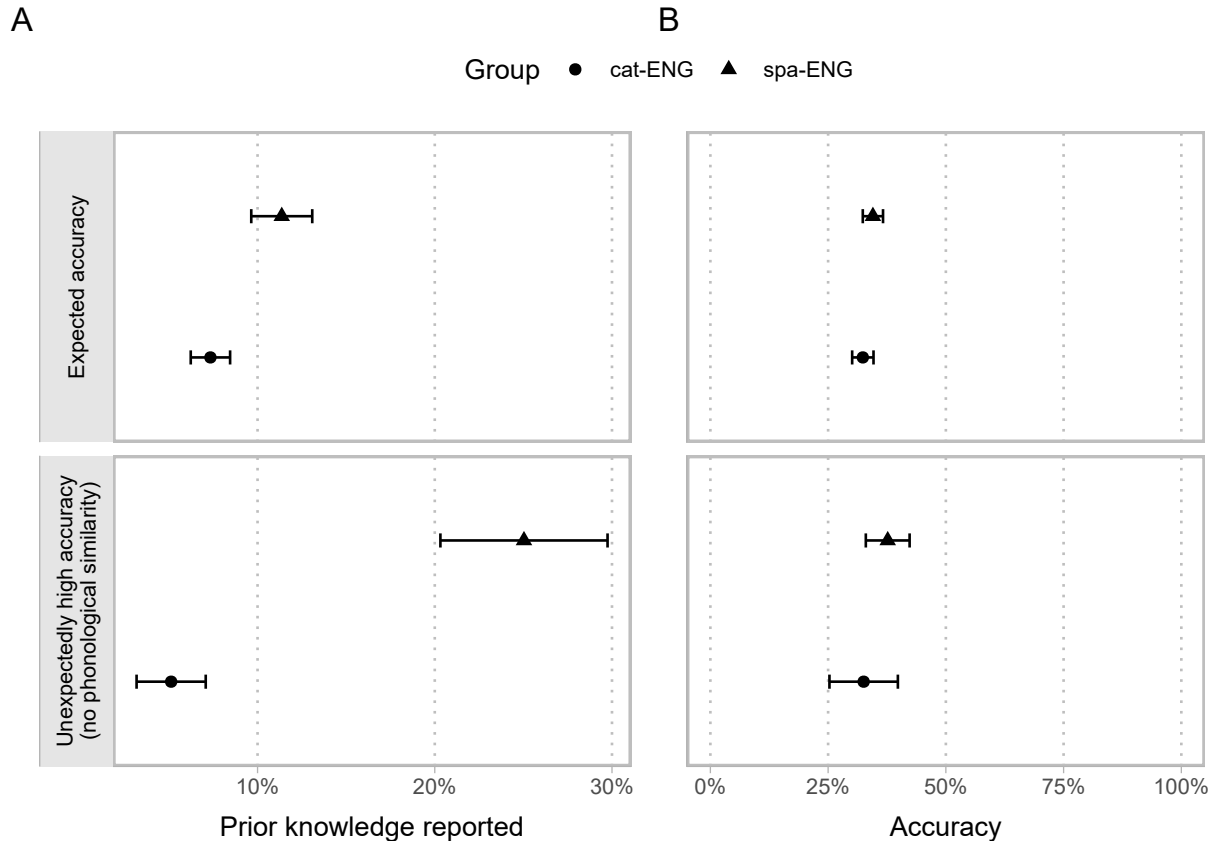
Figure 5 Catalan/Spanish prior word knowledge as reported by English native participants in Experiment 3. (A) Average proportion of participants that reported prior knowledge for words with surprisingly high accuracy (no phonological similarity with the correct translation, accuracy higher than 10%), and for words with expected accuracy (low similarity-low accuracy, or high similarity-high accuracy). (B) Average accuracy for words with expected and unexpected accuracy.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Grantee Submission*, *76*(1), 1–32.

Chen, H.-C., & Leung, Y.-S. (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(2), 316.

Christoffels, I. K., De Groot, A. M., & Kroll, J. F. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language*, *54*(3), 324–345.

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.

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of english phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America*, *116*(6), 3668–3678.

De Groot, A. M., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, *50*(1), 1–56.

Desroches, A. S., Friesen, D. C., Teles, M., Korade, C. A., & Forest, E. W. (2022). The dynamics of spoken word recognition in bilinguals. *Bilingualism: Language and Cognition*, *25*(4), 705–710.

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.
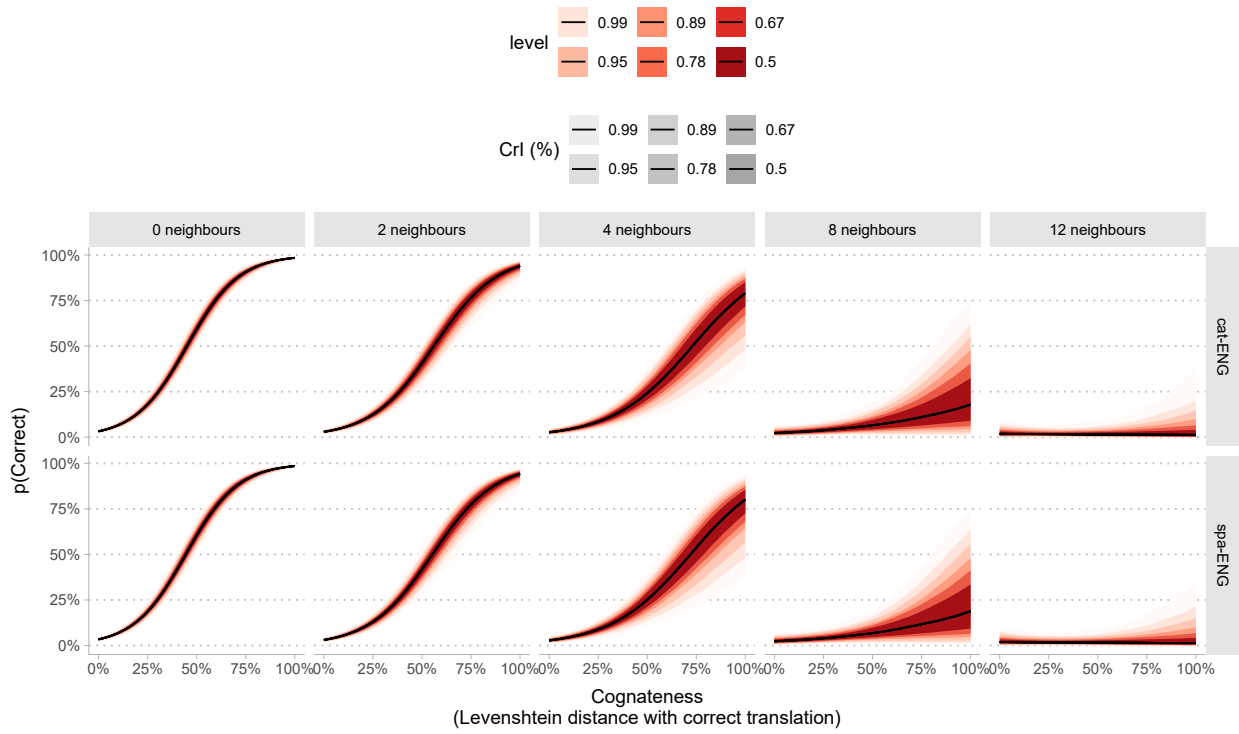
Figure 6Posterior model-predicted mean accuracy in Experiment 2. Predictions were generated from 4,000 posterior samples, extracted for different values of CLPN (0, 2, 4, 8, 12) and Cognateness (1-100). Predictions are plotted separately for English participants translating Catalan words, and for English participants translating Spanish words. Lines indicate mean predictions, and intervals indicate 95%, 89%, 78%, 67%, and 50% credible intervals (CrI).
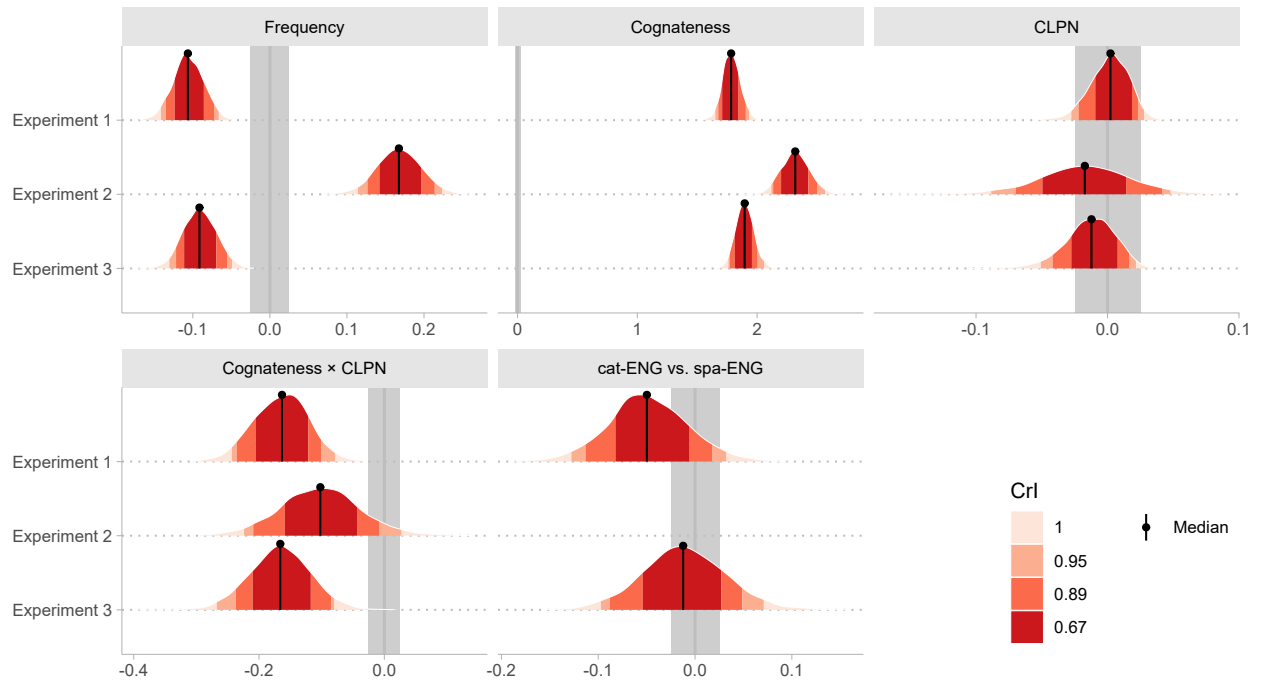

Figure 7

Dufour, S., & Peereman, R. (2003). Lexical competition in phonological priming: Assessing the role of phonological match and mismatch lengths between primes and targets. *Memory & Cognition*, *31*, 1271–1283.

Elias, M., & Degani, T. (2022). Cross-language interactions during novel word learning: The contribution of form similarity and participant characteristics. *Bilingualism: Language and Cognition*, 1–18.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, *28*(5), 501–518.

Groot, A. M. de. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5), 1001.

Groot, A. M. de, Dannenburg, L., & Vanhell, J. G. (1994). Forward and backward word translation by bilinguals. *Journal of Memory and Language*, *33*(5), 600–629.

Hamburger, M., & Slowiaczek, L. M. (1996). Phonological priming reflects lexical competition. *Psychonomic Bulletin & Review*, *3*(4), 520–525.

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.

Kruschke, J. K., & Liddell, T. M. (2018). The bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a bayesian perspective. *Psychonomic Bulletin & Review*, *25*, 178–206.

Lecumberri, M. L. G., Cooke, M., & Cutler, A. (2010). Non-native speech perception in adverse conditions: A review. *Speech Communication*, *52*(11-12), 864–886.

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

Lotto, L., & De Groot, A. M. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning*, *48*(1), 31–69.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1.

Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1990). *Similarity neighborhoods of spoken words.*

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). *CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities.*

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375.

Meyer, A. S., & Damian, M. F. (2007). Activation of distractor names in the picture-picture interference paradigm. *Memory & Cognition*, *35*(3), 494–503.

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.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203.

Phillips, N. A., Klein, D., Mercier, J., & Boysson, C. de. (2006). ERP measures of auditory word repetition and translation priming in bilinguals. *Brain Research*, *1125*(1), 116–131.

Potter, M. C., So, K.-F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, *23*(1), 23–38.

RCore, T. (2019). *R: A language and environment for statistical computing. R foundation for statistical computing, austria.*

Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in europe as based on levenshtein distance. *Bilingualism: Language and Cognition*, *15*(1), 157–166.

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.

Takata, Y., & Nábělek, A. K. (1990). English consonant recognition in noise and in reverberation by japanese and american listeners. *The Journal of the Acoustical Society of America*, *88*(2), 663–666.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, *104*(30), 12530–12535.

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.

van der Loo, M. P. J. (2014). The stringdist package for approximate string matching. *The R Journal*, *6*, 111–122. https://CRAN.R-project.org/package=stringdist

Van Heuven, W. J., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, *39*(3), 458–483.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for british english. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190.

Von Holzen, K., & Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology*, *113*(4), 569–586.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1–25.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Andras, F., Rivera, M., Bajo, T., Dussias, P. E., & Paolieri, D. (2022). Cognate facilitation effect during auditory comprehension of a second language: A visual world eye-tracking study. *International Journal of Bilingualism*, 13670069211033359.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America*, *109*(2), 775–794.

Broersma, P., & Weenink, D. (2021). *Praat: Doing phonetics by computer [computer program]* (Version 6.1.54). http://www.praat.org/

Bürkner, P.-C. (2017). Brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1), 1–28.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Grantee Submission*, *76*(1), 1–32.

Chen, H.-C., & Leung, Y.-S. (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(2), 316.

Christoffels, I. K., De Groot, A. M., & Kroll, J. F. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language*, *54*(3), 324–345.

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.

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of english phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America*, *116*(6), 3668–3678.

De Groot, A. M., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, *50*(1), 1–56.

Desroches, A. S., Friesen, D. C., Teles, M., Korade, C. A., & Forest, E. W. (2022). The dynamics of spoken word recognition in bilinguals. *Bilingualism: Language and Cognition*, *25*(4), 705–710.

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.

Dufour, S., & Peereman, R. (2003). Lexical competition in phonological priming: Assessing the role of phonological match and mismatch lengths between primes and targets. *Memory & Cognition*, *31*, 1271–1283.

Elias, M., & Degani, T. (2022). Cross-language interactions during novel word learning: The contribution of form similarity and participant characteristics. *Bilingualism: Language and Cognition*, 1–18.

Goldinger, S. D., Luce, P. A., & Pisoni, D. B. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, *28*(5), 501–518.

Groot, A. M. de. (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*(5), 1001.

Groot, A. M. de, Dannenburg, L., & Vanhell, J. G. (1994). Forward and backward word translation by bilinguals. *Journal of Memory and Language*, *33*(5), 600–629.

Hamburger, M., & Slowiaczek, L. M. (1996). Phonological priming reflects lexical competition. *Psychonomic Bulletin & Review*, *3*(4), 520–525.

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.

Kruschke, J. K., & Liddell, T. M. (2018). The bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a bayesian perspective. *Psychonomic Bulletin & Review*, *25*, 178–206.

Lecumberri, M. L. G., Cooke, M., & Cutler, A. (2010). Non-native speech perception in adverse conditions: A review. *Speech Communication*, *52*(11-12), 864–886.

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

Lotto, L., & De Groot, A. M. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning*, *48*(1), 31–69.

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1.

Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1990). *Similarity neighborhoods of spoken words.*

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). *CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities.*

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*(5), 375.

Meyer, A. S., & Damian, M. F. (2007). Activation of distractor names in the picture-picture interference paradigm. *Memory & Cognition*, *35*(3), 494–503.

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.

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*(1), 195–203.

Phillips, N. A., Klein, D., Mercier, J., & Boysson, C. de. (2006). ERP measures of auditory word repetition and translation priming in bilinguals. *Brain Research*, *1125*(1), 116–131.

Potter, M. C., So, K.-F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, *23*(1), 23–38.

RCore, T. (2019). *R: A language and environment for statistical computing. R foundation for statistical computing, austria.*

Schepens, J., Dijkstra, T., & Grootjen, F. (2012). Distributions of cognates in europe as based on levenshtein distance. *Bilingualism: Language and Cognition*, *15*(1), 157–166.

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.

Takata, Y., & Nábělek, A. K. (1990). English consonant recognition in noise and in reverberation by japanese and american listeners. *The Journal of the Acoustical Society of America*, *88*(2), 663–666.

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences*, *104*(30), 12530–12535.

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.

van der Loo, M. P. J. (2014). The stringdist package for approximate string matching. *The R Journal*, *6*, 111–122. https://CRAN.R-project.org/package=stringdist

Van Heuven, W. J., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, *39*(3), 458–483.

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for british english. *Quarterly Journal of Experimental Psychology*, *67*(6), 1176–1190.

Von Holzen, K., & Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology*, *113*(4), 569–586.

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language*, *50*(1), 1–25.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686