The role of cognateness in recognition of spoken words in a foreign language

A Preprint

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

July 22, 2024

Abstract

Understanding spoken words in a non-native language is easier when 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. 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 comprehension to 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 on 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

Learning a new language is a non-trivial task, a large part of which is acquiring a new set of word-forms to correspond to existing semantic representations in their lexicon. Imagine that a Spanish listener hears English for the first time and encounters the word /ˈdɔː/ (*door*). Their knowledge of the phonological form in their native language is of little help: /ˈpweɾ.ta/ and /ˈdɔː/ do not share any phonemes. It would come as a surprise if the listener was able to translate the English word successfully without having first learnt the word. However, there are some elements in language that may help the learner in acquiring words in a new language. Many languages share some degree of similarity. This may be due to typological closeness and/or socio-historical events involving the speakers of these languages (e.g., migration, social contact). For example, Romance languages such as Spanish and Catalan share many form-similar translation equivalents (also known as cognates) (Schepens et al., 2012), as in the case of *puerta* and *porta* (*door* in Spanish and Catalan, respectively). Given no prior knowledge of either language, a Spanish native speaker is likely to be much more successful at correctly translating Catalan *porta* than English *door*, due to the phonological and orthographical similarity of the former to the equivalent in their native language. Second language learners have been found to learn cognates more easily than non-cognates (Hall, 2002; Otwinowska & Szewczyk, 2019). This type of cross-linguistic similarity has also been extremely informative for the study of bilingual lexical processing: words that share high form-similarity with their translations are processed faster and more accurately than words sharing lower phonological similarity. This phenomenon has provided strong evidence in favour of the language-non selective hypothesis of bilingual lexical access, which states that bilinguals activate lexical representations in both languages, even during monolingual situations. This facilitation effect has been reported for comprehension (Dijkstra et al., 2010; Midgley et al., 2011; Thierry & Wu, 2007), production (Costa et al., 2000), learning (De Groot & Keijzer, 2000; Elias & Degani, 2022; Lotto & De Groot, 1998; Valente et al., 2018), and translation (Christoffels et al., 2006).

Event counting on this facilitation effect of form-similarity, 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). One of potential source of this increased difficulty is the partial 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 listening to French for the first time. This listener may encounter the word /ˈpɔʁt/ (*porte*), translation of door in French. When comparing the phonemic transcription to the transcription of the Spanish word /ˈpweɾ.ta/ (*puerta*), one may be led to infer that these two words do not have much overlap. But one thing to note is that the voiced uvular fricative consonant /ʁ/ and the open-mid back rounded vowel /ɔ/, which do not exist in Spanish as phonemic categories, can be perceived as allophonic variations the native (Spanish) phonemes /ɾ/ and /o/. A perceptual assimilation model has been proposed to account for how speakers may integrate non-native phonemes into their native phonemic categories based on goodness-of-fit (Best et al., 1988, 2001). Although phonetic mismatch between the non-native word and its translation can have a noticeable toll on comprehension Cutler et al. (2004), the Spanish monolingual may still be able to activate the phonological representation of /ˈpweɾ.ta/ (*puerta*) from hearing the word-form /ˈpɔʁt/ (*porte*), ultimately allowing them to access the right semantic representation.

While phonological similarity across translation equivalents can help with speech recognition, phonological similarity between words with different meanings can hinder the recognition process. In word retrieval, the presence of phonological neighbours (word-forms that differ from each other in one phoneme) may interfere with the activation and selection of the appropriate lexical representation in the native language. Words with denser phonological neighbourhoods are recognised more slowly and less accurately than words with sparser phonological neighbourhoods (Dufour & Peereman, 2003; Goldinger et al., 1989; Hamburger & Slowiaczek, 1996; Luce et al., 1990; Luce & Pisoni, 1998). This is especially true if such neighbours share higher phonological similarity with the presented word, or are lexically more frequent than the target translation (Luce & Pisoni, 1998). The inhibitory effect of phonological neighbourhoods has also been reported across languages. For instance, Weber & Cutler (2004) observed that bilinguals displayed lexical competition effects from words in their native language when performing a word-referent matching task in their L2. The assimilation of non-native phonemes into native categories (Best et al., 1988, 2001) was used to explain how lexical competitors in the native language can be activated even when the target words contain non-native phonemes. In the case of non-native speech recognition, the presence of cross-linguistic pairs which are phonologically similar but differ in meaning (e.g., false friends) may act as distractors during lexical access, obstructing the selection of the appropriate target translation in the listener's lexicon. For instance, Otwinowska & Szewczyk (2019) reported that false friends were disadvantaged relative to non-cognates by Polish second language learners, in contrast to cognates which were known better. It is therefore important to investigate the joint effect of both cognates and false friends when investigating the effect cross-linguistic phonological similarity has on word recognition in a foreign language. This could shed light on available strategies and challenges associated with the early stages of second language acquisition.

In the present study, we investigate the extent to which monolinguals can exploit phonological similarity between translation equivalents when listening to words in an unfamiliar language, and how it interacts with competition from phonological neighbours in the native language. We used a translation elicitation task in which participants listened to words from an unfamiliar language, and then tried to guess their translation in their native language. We will henceforth refer to the auditory-presented words heard by participants on each trial as presented words, and the correct translation for the presented words as target translations. We explored listeners' reliance on phonological similarity to succeed in the task by manipulating the amount of phonological similarity between the presented words and their target translations. Since participants were unfamiliar with the presented language, they could only use phonological similarity between the presented language and their native language to successfully translate the words. We therefore predicted that participants' performance should increase when the translation pairs are phonologically more similar. We also predicted that there is a minimum threshold of phonological similarity to be sufficient for translation.

We further investigated the effect of phonological competitors. Even if the presented word and its target translation share high phonological similarity, participants may still provide incorrect translations if the presented word also shares high phonological similarity with other words in the native language. The effect is predicted to be especially strong if any of the competitors is more lexically frequent than the target translation. We hypothesised that the size of the facilitatory role of phonological similarity would be inversely proportional to the amount of higher-frequency phonological neighbours of the presented word in the target language. For instance, one could expect English participants to incorrectly translate the Spanish word *botón* as *bottom* instead of as its correct translation *button*, due to combined effect of the close phonological similarity between *botón* and *bottom* along with the high lexical frequency of *bottom* relative to *button*. To test this prediction, we developed a lexical frequency-dependent measure of cross-language phonological neighbourhood density, in which a neighbour is counted only if it is higher frequency and is one phoneme apart from the presented word. If the phonological neighbourhood density of the target translation affects participants' performance negatively, this would suggest that competitors in the native language affect recognition of non-native words during recognition of foreign speech.We conducted a series of three experiments to test these predictions.

In Experiment 1, we collected data from two groups of British English natives living in the United Kingdom. One group was presented with Catalan words, the other with Spanish words. We examined the extent to which participants were able to use the phonological similarity between the presented word (in Catalan or Spanish) and its target translation to provide accurate responses. In Experiment 2, we tested a group of (European) Spanish natives in the same task, who were presented with Catalan words. Catalan and Spanish are both Romance languages whose close typological distance is reflected in the fact that they share many cognates, where English is a Germanic language that shares considerably fewer cognates with Catalan and Spanish. By testing participants translating words from typologically close or distant languages, we expected to widen the range of the phonological similarity scores of the translation pairs involved in the experimental task, therefore allowing us to explore potential cross-language differences in participant's performance. One unexpected finding was that participants in Experiments 1 and 2 were surprisingly good at translating a subset of words which had low phonological similarity with their correct translation. We were concerned that this may be caused by some prior knowledge of specific words by our participants, as some Spanish words are commonly seen in media or product labels, making them familiar even to monolingual speakers of English. In Experiment 3, we collected additional data from a new group of British English natives. The design was closely modelled after Experiment 1, except that after providing their response in each trial, participants reported whether they had previous knowledge of the presented word. In the final section of this paper, we present analyses on the joint 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 (*Similarity*), 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).

We calculated *Similarity*, our main predictor of interest, by computing the Levenshtein similarity between the X-SAMPA transcriptions of each pair of translations using the `stringdist` R package (van der Loo, 2014). The Levenshtein distance computes the edit distance between two character strings—in this case, phoneme transcriptions—by counting the number of additions, deletions, and substitutions necessary to make both strings identical (Levenshtein, 1966). We divided this edit distance by the number of characters included in the longest X-SAMPA transcription of the translation pair. This results in a proportion score, in which values closer to zero correspond to lower Levenshtein distances between phonological transcriptions (i.e., higher similarity), and values closer to 1 correspond to higher Levenshtein distances (i.e., lower similarity). This transformation also help account for the fact that the Levenshtein distance tends to increase with the length of the transcriptions. For interpretability, we subtracted this proportion from one, so that values closer to one correspond to higher similarity between phonological transcriptions, and values closer to zero correspond to lower similarity between phonological transcriptions. For example, the *table* (`teIb@l`)-*mesa* (`mesa`) translation pair had a 17% similarity, while the *train* (`trEIn`)-*tren* (`tREn`) translation pair had a 60% similarity. Table 2 summarises the lexical frequency, phonological neighbourhood density and phonological overlap of the words included in the Catalan and the Spanish lists.

For each Catalan and Spanish word, we calculated the number of *CLPN* by counting the number of English words with same or higher lexical frequency, and whose phonological transcription (in X-SAMPA format) different in up to one phoneme from that of the presented Catalan or Spanish word. Lexical frequencies and phonological transcriptions were extracted from the multilingual database CLEARPOND (Marian et al., 2012)[1].

Table 2: Stimuli properties.

	Frequency		Similarity		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

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

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

[1] [Phonological transcriptions in CLEARPOND were generated from eSPEAK, http://espeak.sourceforge.net/]

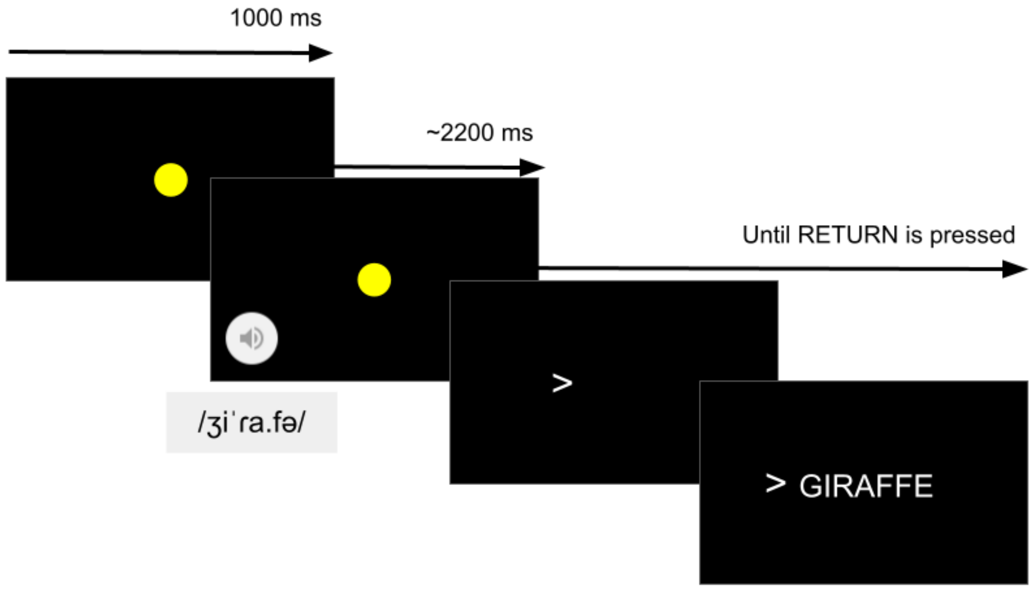

Figure 1: Schematic representation of a trial in the experimental task

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 (5,694 in Catalan, 3,458 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*, *Similarity*, and *CLPN*, and the two-way interaction between *Similarity* 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.

$$
\begin{aligned}
& \textbf{Likelihood} \\
y_i \sim & \text{Bernoulli}(p_i) \\
\\
& \textbf{Parameters} \\
\text{Logit}(p_i) = & \beta_{0[p,w]} + \beta_{1[p]}\text{Frequency}_i + \beta_{2[p]}\text{PTHN}_i + \beta_{3[p]}\text{Similarity}_i + \beta_{4[p]}(\text{PTHN}_i \times \text{Similarity}_i) \\
\beta_{0-6[p,w]} \sim & \mathcal{N}(\mu_{\beta_j}, \sigma_{\beta_j}), \text{ for participant } p \text{ in } 1, ..., P \text{ and word } w \text{ in } 1, ..., W \\
\beta_{1-6[p]} \sim & \mathcal{N}(\mu_{\beta_j}, \sigma_{\beta_j}), \text{ for participant } p \text{ in } 1, ..., P \\
\\
& \textbf{Prior} \\
\mu_{\beta_{p,w}} \sim & \mathcal{N}(0, 0.1) \\
\sigma_{\beta_p}, \sigma_{\beta_w} \sim & \text{HalfCauchy}(0, 0.1) \\
\rho_p, \rho_w \sim & \text{LKJCorr}(8)
\end{aligned}
$$

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 (R Core Team, 2013). 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 4 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 *Similarity* and *CLPN* (β = -0.653, 95% CrI = [-0.973, -0.313], p(ROPE) = 0). As anticipated, participants' performance benefited from an increase in *Similarity* (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 2 illustrates the posterior of the average predictions of the model for words with different values of *Similarity* and *CLPN*. Figure 6 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
Mean	—	20.33	6.62	3.62	—	9,010.08	—	112.22	—

2.3 Discussion

In the present experiment, we investigated the extent to which the 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 phonological similarity 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.

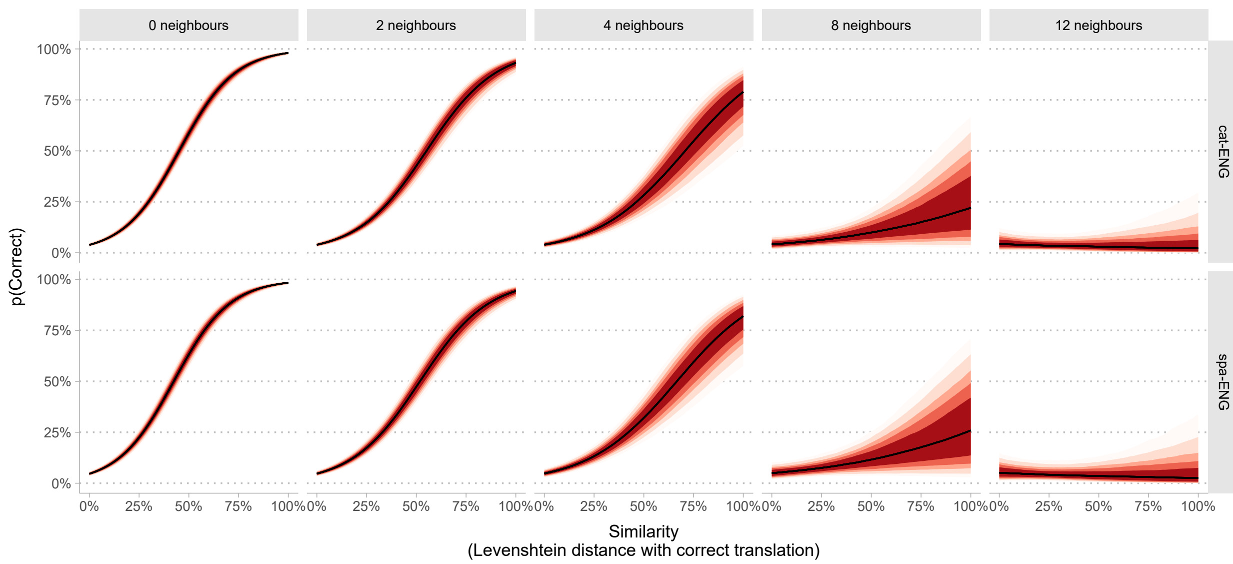

Figure 2: Posterior model-predicted mean accuracy in Experiment 1. Predictions were generated from 4,000 posterior samples, extracted for different values of CLPN (0, 2, 4, 8, 12) and Similarity (1-100). Predictions are plotted separately for English participants translating Catalan words, and for English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

3 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 *Similarity* and *CLPN* (β = -0.409, 95% CrI = [-0.895, 0.111], p(ROPE) = 0.081). As anticipated, participants' performance benefited from an increase in *Similarity* (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 2 illustrates the posterior of the average predictions of the model for words with different values of *Similarity* and *CLPN*.

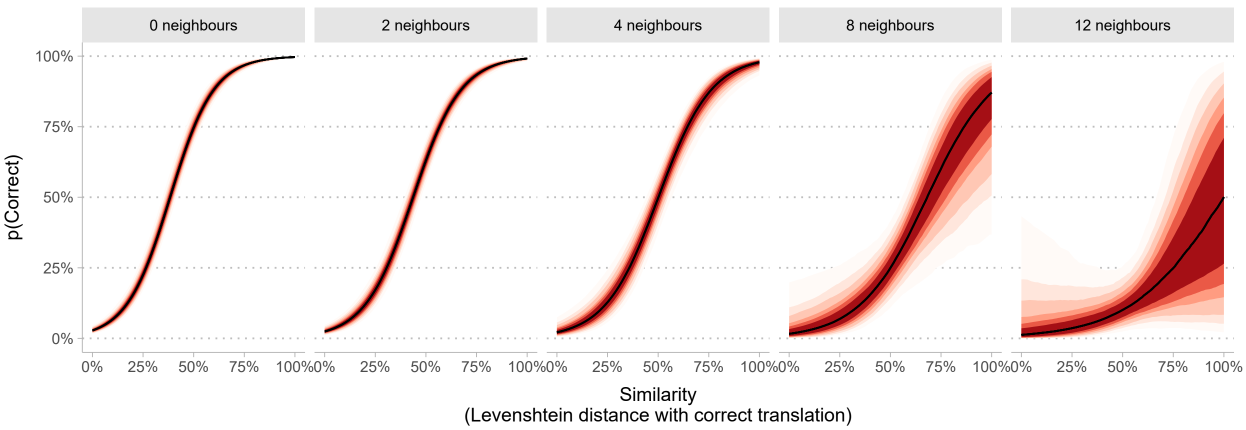

Figure 3: Posterior model-predicted mean accuracy in Experiment 1. Predictions were generated from 4,000 posterior samples, extracted for different values of CLPN (0, 2, 4, 8, 12) and Similarity (1-100). Predictions are plotted separately for English participants translating Catalan words, and for English 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 the results in Experiment 1, the positive impact of phonological similarity on participants' performance in Experiment 2 was more resilient to the interference of phonological neighbours. English natives in Experiment 1 provided significantly less accuracy responses when more than four phonological neighbors were present (even when translating high-similarity words), compared to when only one or neighbour were present. Spanish participants in Experiment 2 benefited from phonological similarity, even when eight neighbours were present. Spanish participants' performance declined after 8 neighbours, and was evident at 12 neighbours. Overall, this suggests that participants in Experiment 2, who were natives of a typologically similar language (Spanish) to the presented language (Catalan) benefited more strongly from phonological similarity than participants in Experiment 1, who were natives of typologically less similar language (English) to the presented language (Catalan, Spanish).

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

Table 5: List 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 /paŋtalon/ - 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

It is likely that participants had prior knowledge of these words despite having reported little to no familiarity 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 4). 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.

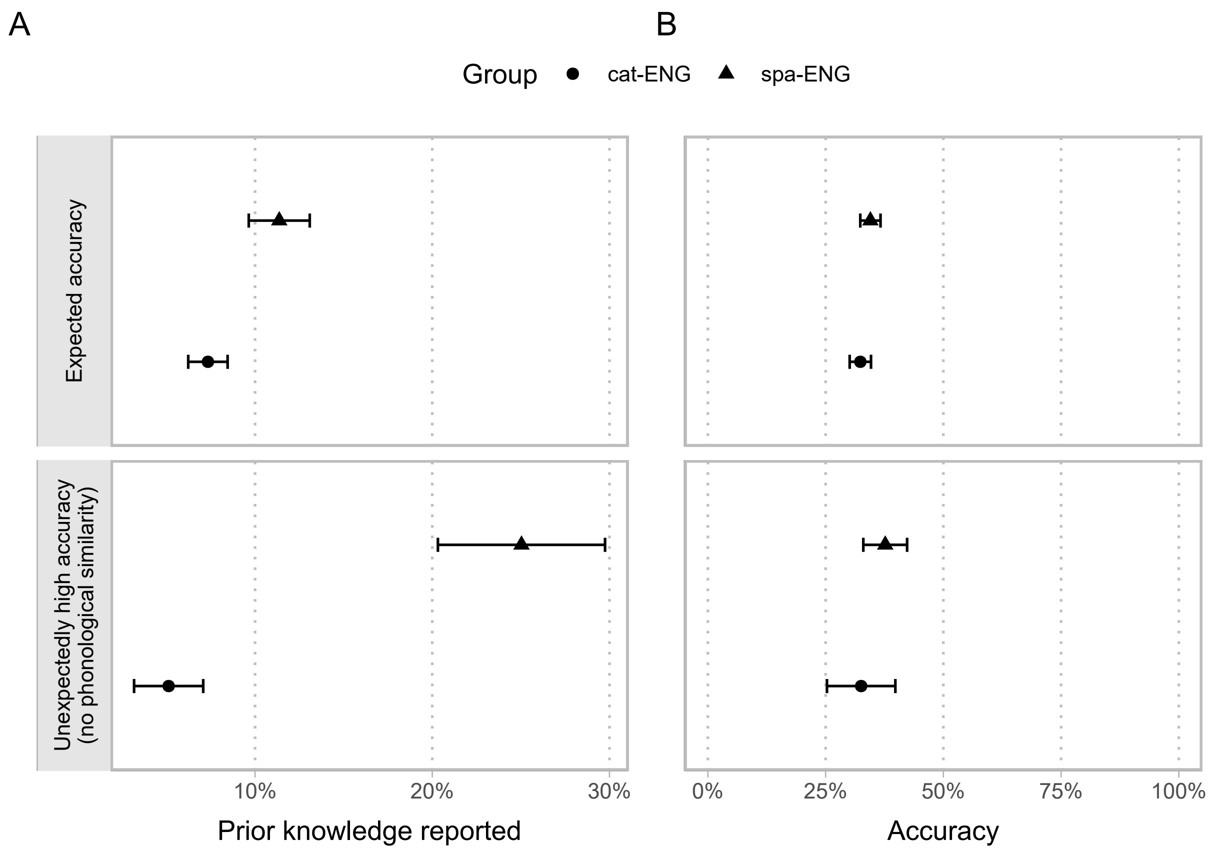

Figure 4: Catalan/Spanish prior word knowledge as reported by English native participants in Experiment 3. (A) Average proportion of participants that reported prior knowledge for words with surprisingly high accuracy (no phonological similarity with the correct translation, accuracy higher than 10%), 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.

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 *Similarity* and *CLPN* (β = -0.665, 95% CrI = [-1.068, -0.32], p(ROPE) = 0.002). As anticipated, participants' performance benefited from an increase in *Similarity* (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 2 illustrates the posterior of the average predictions of the model for words with different values of *Similarity* and *CLPN*.

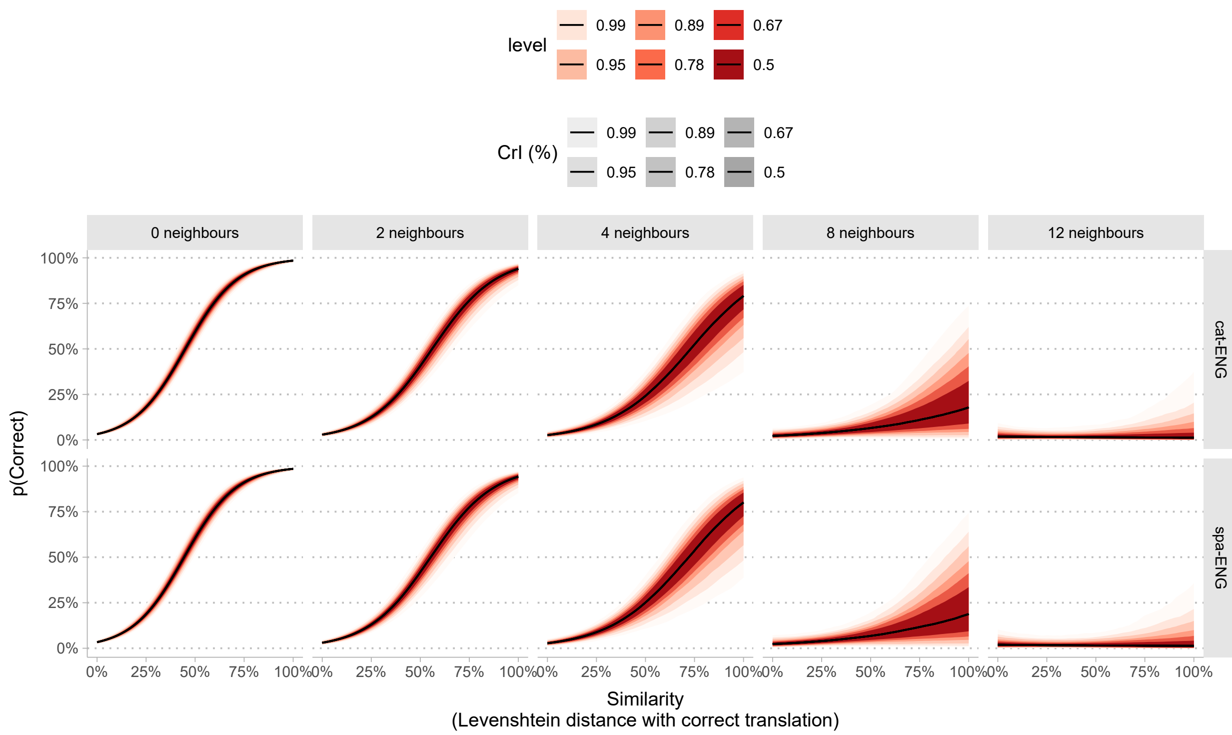

Figure 5: Posterior model-predicted mean accuracy in Experiment 2. Predictions were generated from 4,000 posterior samples, extracted for different values of CLPN (0, 2, 4, 8, 12) and Similarity (1-100). Predictions are plotted separately for English participants translating Catalan words, and for English participants translating Spanish words. Lines indicate mean predictions, and intervals indicate 95%, 89%, 78%, 67%, and 50% credible intervals (CrI).

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 6 summarizes the posterior distribution of the regression coefficients of the models in Experiments 1 to 3.

6 General discussion

7 References

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.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by english-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 345.

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., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1), 1–32.

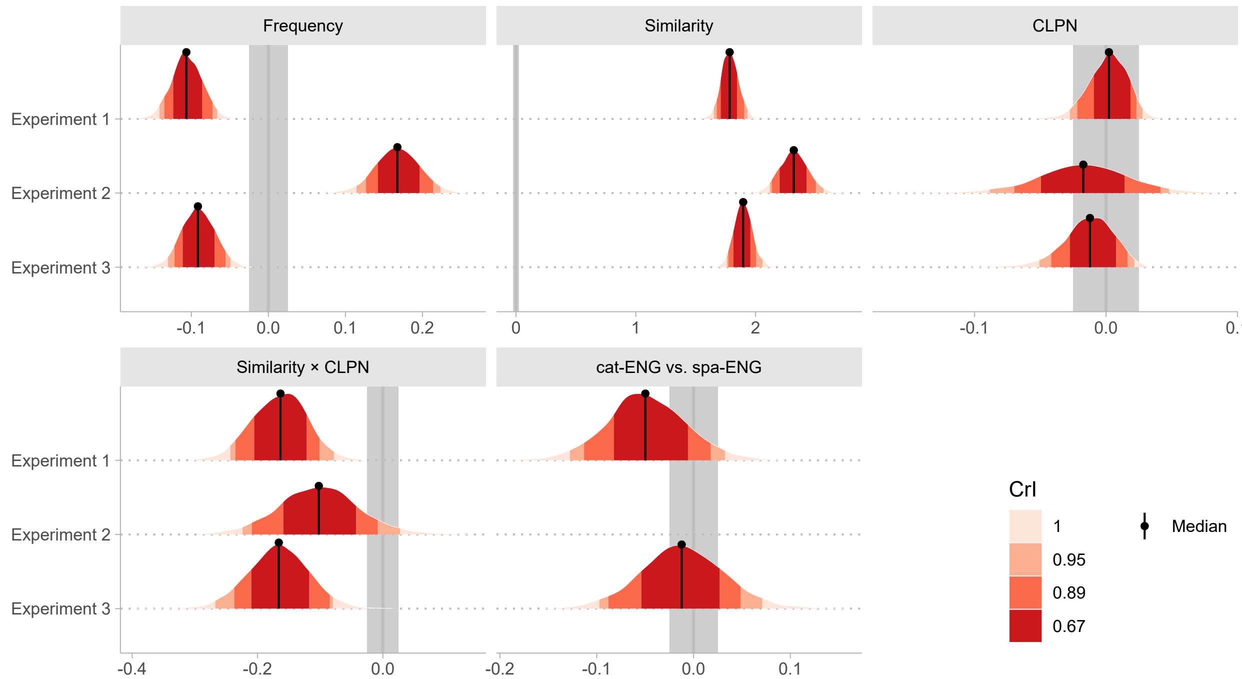

Figure 6

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.

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.

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. https://doi.org/10.1016/0749-596X(89)90009-0

Hall, C. J. (2002). *The automatic cognate form assumption: Evidence for the parasitic model of vocabulary development*.

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

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. (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.*

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.

Otwinowska, A., & Szewczyk, J. M. (2019). The more similar the better? Factors in learning cognates, false cognates and non-cognate words. *International Journal of Bilingual Education and Bilingualism.*

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. https://doi.org/10.3758/s13428-018-01193-y

R Core Team. (2013). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing. http://www.R-project.org/

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.

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*(1), 111–122. https://CRAN.R-project.org/package=stringdist

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.

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

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.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by english-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 345.

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., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1), 1–32.

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.

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.

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. https://doi.org/10.1016/0749-596X(89)90009-0

Hall, C. J. (2002). *The automatic cognate form assumption: Evidence for the parasitic model of vocabulary development*.

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

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. (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*.

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.

Otwinowska, A., & Szewczyk, J. M. (2019). The more similar the better? Factors in learning cognates, false cognates and non-cognate words. *International Journal of Bilingual Education and Bilingualism*.

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. https://doi.org/10.3758/s13428-018-01193-y

R Core Team. (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. http://www.R-project.org/

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.

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*(1), 111–122. https://CRAN.R-project.org/package=stringdist

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.

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