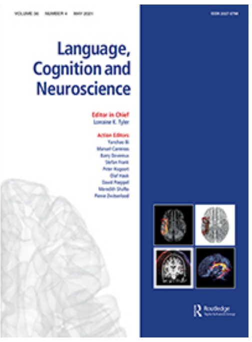

Language, Cognition and Neuroscience

ISSN: 2327-3798 (Print) 2327-3801 (Online) Journal homepage: www.tandfonline.com/journals/plcp21

Lexical activation in late bilinguals: effects of phonological neighbourhood on spoken word production

Solène Hameau, Britta Biedermann & Lyndsey Nickels

To cite this article: Solène Hameau, Britta Biedermann & Lyndsey Nickels (2021) Lexical activation in late bilinguals: effects of phonological neighbourhood on spoken word production, Language, Cognition and Neuroscience, 36:4, 517-534, DOI: 10.1080/23273798.2020.1863438

To link to this article: https://doi.org/10.1080/23273798.2020.1863438

View supplementary material

Published online: 07 Jan 2021.

Submit your article to this journal

Article views: 710

View related articles

View Crossmark data

Citing articles: 4 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=plcp21

LANGUAGE, COGNITION AND NEUROSCIENCE
2021, VOL. 36, NO. 4, 517–534
https://doi.org/10.1080/23273798.2020.1863438

REGULAR ARTICLE

Lexical activation in late bilinguals: effects of phonological neighbourhood on spoken word production

Solène Hameau [a,b], Britta Biedermann [c] and Lyndsey Nickels [a,b]

[a]Department of Cognitive Science, Macquarie University, Sydney, Australia; [b]Macquarie Centre for Reading, Macquarie University, Sydney, Australia; [c]School of Occupational Therapy, Social Work and Speech Pathology, Curtin University, Perth, Australia

ABSTRACT
This research explores patterns of lexical activation of both languages of a bilingual when producing spoken words. Specifically, it investigates the influence of within- and cross-language measures of phonological neighbourhood, on the English picture naming performance of a diverse group of French (L1) – English (L2) late bilinguals. A novel phonological neighbourhood density measure was used that took into account the phonological and lexical constraints of this sample of late bilinguals. Results showed that both within- and cross-language neighbours influenced picture naming performance and that the length of exposure to L2 and familiarity of the target modulated these effects, with both facilitatory and inhibitory effects observed. Additionally, the previously well-attested cognate facilitation effect decreased with L2 language exposure. These results support interactive models of word production, within a language non-specific view of lexical access, and show the dynamic nature of a bilingual's lexical connections.

ARTICLE HISTORY
Received 8 April 2020
Accepted 1 December 2020

KEYWORDS
Late bilingual; lexical retrieval; phonological neighbourhood density; cross-linguistic influence; L2 exposure; cognate

Introduction

When exposed to a new linguistic environment, individuals who acquire a second language later in life naturally experience changes in their lexicon due to the need to progressively integrate new lexical representations and add a new phonological repertoire. These new representations co-exist with those of the native language, with which they may share some degree of form similarity. Phonological neighbourhood is a common way to characterise form similarity between representations in the lexicon. For a given word, phonological neighbourhood *density* is usually defined as the number of words in the lexicon that differ from that word by one phoneme, either substituted, deleted or added (Luce, 1987). For instance, for a (monolingual) speaker of English, phonological neighbours of the word *cat* include the words *fat*, *cab*, *kit*, *scat*, *at*. For a bilingual speaker, a word's phonological neighbours can come from either of the person's languages. In a late bilingual, as vocabulary expands, the number of phonological neighbours of the second language increases, most rapidly in the case of a more recently acquired language. Cognates are words that share a similar form to their translation equivalent (e.g. *music* and its French translation equivalent *musique*). It has been proposed that cognates can be considered a specific case of cross-language phonological neighbours (e.g. Costa et al., 2005). The research reported here explored the extent to which similar form representations in the lexicon affect the ease with which a bilingual can access and produce spoken words in a given language, and which factors potentially modulate this influence.

In (monolingual) language processing, it is generally assumed that when a given word is activated, phonological neighbours of this word are also activated (e.g. Luce et al., 1990). This is the case for the perception of spoken words (e.g. Garlock et al., 2001; Vitevitch & Luce, 1998, 1999), but some theories of production also hypothesise activation of these phonological neighbours during word production via feedback of activation from the phoneme level to the lexical level (e.g. Chen & Mirman, 2012; Dell & Gordon, 2003, building on the lexical network account by Dell et al., 1997). In bilingual speakers, for whom words similar to a given target can exist in both languages, it is well established that cognate words are faster and more accurately named (e.g. Costa et al., 2000; Costa et al., 2005; Strijkers et al., 2010), and this has been suggested to reflect concurrent activation of the translation equivalent's form in the non-target language (e.g. Costa et al., 2005; but see e.g. Kirsner et al., 1993; van Hell & de Groot, 1998, for alternative interpretations of the origin of the cognate

CONTACT Solène Hameau solene.hameau@mq.edu.au
Supplemental data for this article can be accessed at https://doi.org/10.1080/23273798.2020.1863438.
© 2021 Informa UK Limited, trading as Taylor & Francis Group

facilitation effect). Additionally, the cognate facilitation effect seems to be stronger when naming in a speaker's second acquired language (L2) compared to their first language (L1; e.g. Costa et al., 2000; Starreveld et al., 2014; Strijkers et al., 2010), showing the influence of language dominance on patterns of cross-language activation. Finally, in addition to the cognate facilitation effect that suggests activation of the translation equivalent, there is evidence showing that words from the non-target language that are related in form (but are not necessarily translation equivalents) are activated up to the word form level in word production (e.g. Colomé, 2001). Indeed, it has been argued that there may be feedback between phoneme and word form levels across both languages (e.g. Colomé & Miozzo, 2010; Costa et al., 2006). It could then be the case that both within-language *and* cross-language phonological neighbours play a role in the spoken word production of bilingual speakers.

Effects of phonological neighbourhood density have been attested in the monolingual literature. In the area of spoken word perception, for instance, phonological neighbourhood density seems to have a consistent inhibitory effect (e.g. Luce & Pisoni, 1998) whereas facilitatory effects have been observed on reading aloud (e.g. Mulatti et al., 2006). In spoken word production, there does not seem to be a consensus: the presence and direction of phonological neighbourhood density effects vary across studies and may also vary cross-linguistically (see Sadat et al., 2014, for a review). We have recently found evidence suggesting that the presence and direction of the effects of phonological neighbours might depend on the relative frequency of the target compared to its phonological neighbours (Hameau et al., 2018): Phonological neighbours of higher frequency than the target resulted in inhibitory effects on naming of lower frequency targets, but there was a trend towards facilitatory effects when the target was higher in frequency. Hence, it is also possible that in bilinguals, potential effects of phonological neighbourhood density (within and across bilinguals' languages) are modulated by the frequency of the target word.

In bilingual picture naming, although cognates have been extensively researched, few studies have examined the influence of either within-language or cross-language phonological neighbourhood density. Marian and Blumenfeld (2006) compared the effects of within-language phonological neighbourhood density with two groups of German-English late bilinguals: native speakers of German and native speakers of English. Participants were tested in German, which was either their L1 or their L2. Overall, naming accuracy was better for targets with high neighbourhood density, but there were only effects on latency in participants' second, non-native, language. The authors concluded that words in the non-native, less proficient language were overall lower in frequency and thus more sensitive to phonological neighbourhood effects. This can be taken as evidence that language proficiency, and possibly frequency of the target, modulate effects of phonological neighbourhood density in bilinguals. Cross-language phonological neighbourhood density refers to the number of words in the other language of a bilingual speaker that are neighbours of a given target: for example, French neighbours of English words. In word processing, significant effects have been reported of cross-language orthographic neighbourhood density in visual recognition tasks (e.g. Coltheart et al., 1977; van Heuven et al., 1998). However, in speech production, research is more limited. de Groot et al. (2002) found significant effects of within- and cross-language orthographic neighbourhood density on reading aloud latency in English-Dutch bilinguals. In picture naming, to our knowledge, only one study has looked at the influence of cross-language phonological neighbourhood density: Sadat et al. (2016) examined Spanish picture naming performance in highly proficient (Catalan-Spanish) early bilinguals. Although phonological neighbourhood density was not the primary focus of their study, their predictors of picture naming included both within- and cross-language phonological neighbourhood density. They found an inhibitory effect of within language (Spanish) phonological neighbourhood density on latencies, and no effect on accuracy. There was no effect of cross-language phonological neighbourhood density (number of Catalan neighbours) on response time or on accuracy. To our knowledge, no study has examined cross-language phonological neighbourhood density in late bilinguals.

Calculating phonological neighbours in late bilinguals

Calculating cross-language phonological neighbourhood density seems far from straightforward. For example, Marian and Blumenfeld (2006) explicitly rejected examining its influence, noting that "differences in phonetic features between German and English [preclude] meaningful computations of cross-linguistic phonological neighbourhoods in English" (p. 13–14). Sadat et al. (2016) did not raise any concerns regarding Catalan-Spanish cross-language neighbours, possibly because, although there are some differences, Spanish and Catalan phonological inventories are more similar than those of English and German.

It is well established that, as a consequence of early adaptation to one phonological system, monolingual adults have trouble categorising, discriminating and producing phonemes from other languages when these phonemes are not used to differentiate words of their native language. This is not only the case for naïve listeners of the non-native language (as demonstrated by the Perceptual Assimilation Model: Best, 1995), but also for speakers who are learning a second language, be it by formal instruction or by immersion in a second language (L2) environment (as shown by Best & Tyler, 2007; or Fledge, 1995). In fact, even very proficient bilinguals can fail to differentiate between some non-native phonological contrasts, as shown for instance by Pallier et al. (2001): in an auditory lexical decision experiment, Pallier and colleagues found that, in highly proficient early bilinguals (Catalan-Spanish), a word was primed to the same extent when it was preceded by either an identical word or by a minimal pair from the other language differing by only one non-native phonological contrast. This finding shows that the representation of phonological categories in a non-native language, even in very highly proficient speakers of that second language, is certainly different from that of a native speaker of that language. Although the evidence reviewed here covers phoneme perception, perceptual skill level has been shown to be positively correlated with accuracy in producing L2 phonemes (at least for vowels; e.g. Best & Tyler, 2007). In other words, representations in the input lexicon are likely to be similar to the representations in the output lexicon.

Vitevitch (2012) directly addressed the issue of the calculation of cross-language phonological neighbourhood density that arises as a result of the problems non-native speakers may have with phonological contrasts in their second language. He first calculated English-Spanish cross-language neighbours based on the International Phonetic Alphabet (IPA) transcription of words of both languages, and found very few cross-language neighbours. He subsequently tried to take "perceptual assimilation" into account, relating to Spanish speakers' likely perception of English vowels: Based on a study by García Lecumberri and Cenoz Iragui (1997), Vitevitch replaced all English vowels with the Spanish vowel that was perceived as most similar to get a more plausible estimate, for Spanish speakers, of the English words that are similar to Spanish words and vice versa. He re-calculated the proportion of Spanish neighbours of English words with this new coding, although this proportion was still very small (increasing from about 5% to about 13%). His conclusion was that, since they represent such a small proportion of neighbours, cross-language neighbours should not influence processing to a great extent. However, Vitevitch did not examine this hypothesis empirically. Importantly, it is not only calculation of cross-language neighbours that is affected by languages' different phoneme repertoires, the same problem applies to the calculation of *within*-language neighbours in a second language: Phonemes may not be equally perceived and/or produced by a non-native and a native speaker, leading to differences in what may be a neighbour. Consequently, we attempted to take these factors into account when defining measures of phonological neighbourhood in late bilinguals.

The aim of the research reported here was to explore the influence of phonological neighbours on the picture naming performance of late bilinguals, and specifically effects on English picture naming in a population of late bilinguals[1] who had French as a native language. We examined whether different measures of phonological neighbourhood differed in their ability to predict picture naming behaviour in this population. Several measures were considered: first, translation equivalent similarity (a continuous measure of "cognateness": how "neighbour-like" the translation equivalent is to the target); second, the phonological neighbourhood density of the target in the target language (i.e. English), the target's phonological neighbourhood density in the non-target language (French), and a measure combining both within- and cross-language neighbourhood densities. The phonological neighbourhood density measures were calculated taking into account specificities of the phonological system of the late bilinguals (and hence labelled as "bilingual-specific"). However, to allow for comparison with previous research, these measures were compared to measures of phonological neighbourhood density computed the "traditional" way (without any "late bilingual" adaptation).

A theoretical account that is of relevance for the present study is the weaker links hypothesis (e.g. Gollan et al., 2008). The weaker links hypothesis states that, because bilinguals use words in each of their languages less frequently than do monolinguals of the same languages, their lexical representations in both languages will have been retrieved less frequently, leading to "weaker links" between semantic and phonological representations than monolinguals. What is unique about this account is that it relates effects of bilingual experience to frequency effects. While this can be used to explain the differences between monolinguals and bilinguals, these differences would also be apparent between bilinguals that had different amounts of exposure to their second language. The frequency status of words in a given language will then vary as a function of the extent to which this language has been used. The frequency of individual words in a language is also important when evaluating the impact of usage: Gollan et al. (2008) interpret frequency as impacting resting levels of activation, in a system with thresholds for lexical selection (see also e.g. Morton, 1970). The higher the frequency of a given word the closer it will be to its selection threshold. As higher frequency words will be those used most often, even when a particular language is used less frequently, they will nevertheless be close to threshold. It is the lower frequency words that are suggested to be most affected by the degree of use. This leads to the prediction that the effects of bilingualism should be stronger for lower frequency words, but also that frequency effects should be larger in the non-dominant than in the dominant language. Gollan et al. (2008) examined these predictions with a picture naming task, and did find that frequency effects on response time and accuracy were stronger in bilinguals compared to monolinguals, and strongest in the bilingual's less dominant language. Gollan and colleagues hence demonstrated the influence of bilingualism on patterns of lexical activation.

We propose to extend the reasoning behind the weaker links hypothesis to the activation of phonological neighbours. Words in the more dominant language can be thought of as being potentially more highly activated than words in the less dominant language, because of their higher resting level of activation. If by virtue of interactivity between the lexical and post-lexical levels, co-activated words (phonological neighbours) influence the production of a given target word, it follows that these phonological neighbours will have strongest effects if they belong to the most dominant language, and will also have stronger effects on lower frequency targets. Therefore, following this reasoning, we aimed to determine whether potential effects of (any type of) phonological neighbourhood were modulated by two factors: (1) an individual's length of exposure to English, in line with the different effects observed in L1 vs L2, both for phonological neighbourhood density (e.g. Marian & Blumenfeld, 2006) and for cognates (e.g. Costa et al., 2000), and (2) the frequency of the target word, consistent with the fact that different effects of phonological neighbourhood measures were observed on monolinguals depending on the frequency of the target (Hameau et al., 2018). We hypothesised different effects of target language and non-target language phonological neighbourhood density depending on the length of exposure to English, and on the frequency of the English target.

Materials and methods

Participants

Fifty-one individuals (32 females) with French as a native language and who had lived in an English-speaking country for more than two years were recruited through networks of French speakers in Sydney (e.g. Alliance Française, French Lycée, Facebook groups). This study was approved by the Macquarie University Human Research ethics committee. Participants provided written informed consent and received 15AUD for their participation. One of the participants' data was omitted because of low accuracy (below 55%). All participants grew up in a French monolingual environment, either in France or in a French-speaking region of Belgium, learned English at school for at least four years, moved to an English speaking country after age 16, and were living in Sydney at the time of the experiment. The 50 participants whose data were included had normal or corrected-to-normal vision and no reported history of learning difficulties or language impairment. None spoke a language other than French or English at home. They were aged 18 to 65 (mean = 37.22, SD = 11.11), had spent two to 42 years living in an English-speaking country (mean = 9.96, SD = 8.94) and claimed to spend between five and 60% of their time per week speaking French rather than English (mean = 25, SD = 14.03).

We considered (late) bilingualism as a continuum, and differentiated between bilinguals using an objective measure of overall exposure to the L2: The number of years spent living in an English-speaking environment. Although this measure does not take into account the relative use of English and French in an individual's daily life, we argue that in the case of our adult participant sample, individuals are greatly if not primarily differentiated by the time they spent living in an English-speaking environment. For instance, age and context of acquisition were similar in our population, with all participants having lived in a French-speaking environment and only speaking French for at least the first 16 years of their lives, and all participants having first learned English formally at school in their French-speaking country. At the time of the experiment, all the participants were either studying or working in an environment where some use of English was necessary. In addition, for the purpose of our experiment, a measure of "proficiency" in general, was not appropriate as this would include dimensions that are not relevant here (e.g. grammar, syntax, idioms, pronunciation, mastery of written English). Rather, for our purposes, we required a measure that is most likely to reflect (only) the extent of the speaker's L2 vocabulary. For instance, exposure to L2 has been strongly linked to vocabulary acquisition in bilingual children (e.g. Thordardottir, 2011). The number of years of exposure to English is therefore likely to affect both the individual's vocabulary size and how the coexistence of the two lexicons affects processing.

Materials

Stimuli comprised 386 black and white drawings of nouns representing objects taken from the International Picture Naming Project (IPNP) picture database (Szekely et al., 2004) selected to have a target name that was a single word, and of high name agreement in English monolingual speakers (75% or greater). Following the experiment, items were excluded from the analyses if they had low (less than 50%) Australian[2] or bilingual accuracy (56 items), or because the Australian name for this target was comprised of two words (e.g. *skipping rope*, *spinning top*, *coat hanger*; 10 items). Two additional compound words (*backpack* and *bandaid*) were not present in the CELEX database (Baayen et al., 1993) and were also removed. This led to a final set of 318 items. A list of all stimuli is provided in Appendix A.

Procedure

The experiment was approved as part of a larger project by the Macquarie University Human Research Ethics committee, and participants gave informed written consent prior to completing the experiment. DMDX (Forster & Forster, 2003) was used for presentation of the stimuli. Instructions were presented on the screen and explained further by the examiner: participants were asked to name pictures as quickly and accurately as possible, in English, with a single word. Each trial started with a 200 ms fixation cross, followed by a blank screen for 600 ms, which was then followed by the target picture presented in the centre of the screen for 2000 ms. Recording started upon stimulus presentation and continued for 2000 ms after the picture disappeared. A new trial was initiated 1500 ms after timeout. The onset of a vocal response triggered a voice key, determining response time. There were 10 practice items, followed by the 386 picture stimuli appearing in random order, organised in four blocks of approximately 11 min each, separated by a break. The experiment lasted approximately an hour.

Analysis

Response time and accuracy data were analysed using (generalised) linear mixed effects models, using a range of control and experimental predictors.

Control predictors

Our control predictors included psycholinguistic variables (described in Table 1) that have commonly been found to influence the speed of picture naming: visual complexity, imageability, name agreement, familiarity, log word frequency, word length in phonemes (e.g. Alario et al., 2004; Bonin et al., 2002). We chose not to include age of acquisition as age of acquisition norms from native English speakers are uninformative for second language English speakers. We also added two variables relating to the experiment: trial order and the standardised response time of the preceding trial (Baayen & Milin, 2010).

Experimental predictors

Our experimental predictors were language exposure, phonological edit distance (see below), and a set of measures of phonological neighbourhood density.

Language exposure. Language exposure was defined as the number of years the participant had been exposed to English by living, fully immersed, in an English-speaking country: mean = 9.96 years, SD = 8.94.

Phonological edit distance. Phonological edit distance between translation pairs (Phonological Corpus Tools: Hall et al., 2015) could be considered a continuous

Table 1. Summary of item-related control predictors.

Name	Definition	Mean(SD)
Visual complexity	Size of the digitised stimuli picture files in Kbytes (IPNP, Szekely & Bates, 2000)	16,665 (8911)
Imageability	Human ratings of a word's capacity to arouse nonverbal images (MRC database: Coltheart et al., 1977). Only available for 250 of the 318 items.	593 (33)
Name agreement	The degree to which participants agree on the name of the picture. Based on mean speeded accuracy of 40 young monolingual Australian English speakers (Hameau, 2016) expressed in %	88 (13)
Familiarity	Human ratings of how usual or unusual the concept/object is in an individual's realm of experience on a 7-point scale (Johnston et al., 2010)	5.02 (1.15)
Log word frequency	Log summed spoken and written word form frequency (per million) (CELEX: Baayen et al., 1993)	2.39 (0.71)
Word length (phonemes)	Number of phonemes	4.25 (1.58)
Phonological edit distance	Phonological similarity between the English target and its French translation equivalent. Higher scores indicate less overlap (Phonological Corpus Tools: Hall et al., 2015)	18.52 (10.42)

measure of "cognateness". This measure compares the French translation's phonology and the IPA transcription of the English target word, calculating the number of one symbol changes required to turn the English phonological form to the French phonological form. Changes are here weighted based on featural similarity. For instance, the English noun *saw* (/sɔ/) and its French translation equivalent *scie* (/si/) have a phonological edit distance score of 3, which is obtained as follows: the consonant is identical, but the vowel has three different featural contrasts: Open-mid versus close, back versus front, and rounded versus unrounded, hence the phonological edit distance value of 3. The more featural contrasts between each phoneme of a word and the corresponding phoneme of its translation, the higher the phonological edit distance. We chose this measure of formal overlap between the target word and its translation equivalent, rather than a binary cognate/non cognate distinction, to allow a more fine-grained consideration of phonological similarities between the two languages at play in this experiment.

Phonological neighbourhood density. The bilingual-specific measure of phonological neighbourhood density used in this study was tailored to our bilingual population for three reasons: first, we wished to examine neighbours from both the target language (English) and the non-target language (French); second, in late bilinguals the limited size of the L2 vocabulary reduces the number of L2 neighbours compared to monolingual native speakers of that language, and third, the potential specificities of the phonological representations in our sample of late bilinguals may require an adapted calculation of phonological neighbourhood density.

Our calculation method was developed taking into account the following factors. First, there is clearly no full overlap between vowel categories in English and French, with 20 different vowels in English and 17 in French, including nasal vowels. Take, for example, the English target word *mitt* (/mɪt/). In French, there is only one /i/ vowel (at least in the European French varieties), therefore French speakers might not have two clear and distinct representations for the English phonemes /ɪ/ (as in *mitt*) and /i/ (as in *meat*), and both might be perceived as equivalent (i.e. *mitt* and *meat* will be homophones). If that is the case, then for a French native speaker, the English words *bit* and *meat* only differ by one phoneme and, hence, unlike for English monolinguals, count as phonological neighbours. In the same vein, for the late bilingual, the French word *vite* (/vit/ - meaning "fast") would count as a cross language neighbour of *bit*, although strictly following the one phoneme difference rule based on IPA, they differ by two phonemes.

We used Iverson and Evans' (2007) findings to map English vowels to French vowels for French native speakers. They asked French native speakers to listen to aurally presented English vowels, identifying which French vowel was most similar this English vowel, as well as rating how similar the English vowel was to its French so-defined counterpart (see Table 2). All their participants grew up in France but had started learning English in childhood and were tested in London where they had been living for up to 17 years (we note here that this sample is similar to that of the present study in terms of the type and amount of exposure to English). We used these English-French vowel mappings to build on Vitevitch (2012) and develop "French-speaker" transcriptions of English targets: English vowels that were paired to a different French vowel with a similarity rating of 70% or more, were replaced by that vowel. For instance, IPA transcription for the word *bit* (/bɪt/) was recoded as /bit/ as French speakers rated the English vowel /ɪ/ 87% similar to the French vowel /i/.

Phonological neighbourhood densities of target words were calculated using the program CLEARPOND (Marian et al., 2012), with the addition of the neighbours of the "French-speaker" transcription of these words. We then adjusted the pool of potential neighbours, retaining only words that were likely to be present in the majority of our French native speakers' English lexicon, by estimating a frequency cut-off: 27,000 words is the average number of words known by a native speaker of English aged 30 (levels of vocabulary can go up to 40,000). For second language learners after 3–4 years spent in an English-speaking country, the number of words is estimated to 13000 words (so 32 to 48% of a native speaker's vocabulary) (Estimates of vocabulary sizes for both monolinguals and bilinguals were taken from the "Test your vocab" (2013) website: http://testyourvocab.com). Words in the CELEX (Baayen et al., 1993) database

Table 2. English vowels, the most similar corresponding French vowel, and the average similarity rating (percentage) between the English and the French vowels (adapted from Iverson & Evans, 2007).

English vowel	Closest French vowel	Average similarity rating
i	i	.86
ɪ	i	.87
eɪ	ɛ	.70
ɛ	ɛ	.86
a	a	.91
aɪ	a	.57
aʊ	a	.47
ɑ	ɑ	.80
ɒ	ɔ	.83
ɔ	o	.84
əʊ	o	.74
ɜ	ø	.80
ʌ	ø	.85
u	u	.82

most likely reflect the highest level of vocabulary a native speaker of English can achieve. Hence, we estimated that the 32% of words with the highest CELEX frequency were those most likely to be in our bilingual participants' English lexicons. These words correspond to words with a CELEX log combined (sum of spoken and written) word-form frequency greater than 3. French neighbours of the English targets were also obtained using CLEARPOND (Marian et al., 2012)[3] and using the same "French-speaker" transcription in addition to the English form of each target.

We computed several measures of phonological neighbourhood density (PND). These were both measures that used our bilingual-specific approach, and, for better comparison with previous research, traditionally computed measures that do not account for phonological assimilation and frequency in bilinguals, and consist of the raw output from CLEARPOND (Marian et al., 2012):

- English PND, bilingual-specific (mean = 11.70, SD = 12.08).
- English PND, traditionally computed English PND tradit (mean = 15.08, SD = 15.42).
- French PND, bilingual-specific: number of neighbours of the English target in French (mean = 13.02, SD = 21.27).
- French PND, traditionally computed (French PND tradit.; mean = 5.15, SD = 12.77).
- Combined English and French PND (Combined PND), bilingual-specific: Total number of neighbours of the English target in both English and French (sum of PND and French PND) (mean = 24.73, SD = 31.03).
- Combined PND, traditionally computed (Combined PND tradit., mean = 20.95, SD = 24.55).

The measures that took into account late bilinguals characteristics differed significantly from the "traditional" PND measures computed using the raw output from CLEARPOND (English PND measures: $t(600) = 3.73$, $p < .001$; French PND measures: $t(632) = 6.65$, $p < .001$; Combined PND measures: $t(462) = 6.03$, $p < .001$), however, they were strongly correlated (for English PND: $r = .948$, $p < .001$; French PND: $r = .662$, $p < .001$; Combined PND: $r = .865$, $p < .001$).

Analysis procedure

Vocal response latencies were manually adjusted for response onset, in order to correct for instances when the voice trigger was activated by irrelevant noise, using CheckVocal (Protopapas, 2007). A response was coded as correct when it corresponded exactly to the target word, with no fluency error. For example, *refrigerator* for *fridge* or *bathtub* for *bath* were discarded. Incorrect responses included fluency errors, synonyms, visual errors, other error types that did not fit in these categories, and no responses (or timeout).

All predictors were standardised to be on the same scale, so that the interpretation of intercepts was not affected by interaction terms between the variables at play (Schielzeth, 2010). Inspection of the pairwise correlations between all predictors revealed that several of the predictors were significantly correlated, especially and unsurprisingly most of the PND variables (see correlation table in Appendix B). Hence, we had potential multicollinearity between our PND measures (correlation coefficients of .8 or higher: Hutcheson & Sofroniou, 1999). To remediate this, we ran different models for each of the different PND measures. Although this procedure did not allow us to determine the effect of a PND measure "over and above" the effects of the other PND measures, by comparing the fit of respective models we could nevertheless evaluate and compare their predictive power. In addition, we monitored multicollinearity by computing variance inflation factors (VIFs) with values above 2.5 being the sign of potentially problematic multicollinearity (Allison, 2012; throughout the whole set of analyses, all VIFs were well below that threshold).

In addition to the main effects of PND measures, we were interested in testing a number of interactions. Following Hameau et al.'s (2018) results, we examined the interaction between PND measures and the frequency of the target. However, given that frequency of the English words was not felt to be an appropriate measure of "true" frequency for our late bilinguals participants, we instead investigated the interaction between PND measures and familiarity, which correlated strongly with log frequency in our item set ($r = .494$, $p < .001$). It has been shown that familiarity ratings, although closely related to frequency, may be more accurate, especially for words in the lower frequency range (e.g. in lexical decision: Gernsbacher, 1984; in picture naming in aphasia: Nickels & Howard, 1995), and recent research suggests that measures of frequency based on ratings rather than objective measures of frequency have a better predictive power in L2 processing (Chen & Dong, 2019). Our familiarity measure, based on ratings, may be a better fit for our L2 naming data compared to a corpus-based frequency measure, and was therefore preferred to frequency for testing this interaction.

We tested two further interactions (as previously explained) between the length of exposure to English and, on the one hand, PND measures, and on the other hand, phonological edit distance.

We used (generalised) linear mixed effect modelling to assess the specific influence on response time and on accuracy of each phonological neighbourhood value and the continuous translation similarity measure, taking into account the variability inherent to different participants and items. Data were analysed using R (R Development Core Team, 2011) using the lme4 software package (Bates et al., 2014), and lmerTest (Kuznetsova et al., 2017) was used to retrieve *p*-values. Interaction plots were created using the packages effects (Fox & Weisberg, 2019) and ggplot2 (Wickham, 2016). We ran two series of models, first on response time for correct responses, and second on accuracy. Because of generally high correlations between PND measures (see Appendix B), separate models were run for each specific PND measure. How a PND measure contributed to the model fit compared to other PND measures, was assessed by comparing the Akaike Information Criterion of respective models.

Results

The different response time and accuracy models showed remarkably similar effects of the PND measures of interest. An overview of the results is provided in Table 3. In order to avoid repetition in the text, because of the similarity between the results, we describe below the results for the first tested model (involving our late bilingual-specific English PND measure) only, first for response time and then for accuracy. Full model summaries for all the analyses can be found in Appendices C and D.

Response time analyses – bilingual-specific English PND

Incorrect responses were removed from response time (RT) analyses, resulting in 12,198 observations for analysis. For this analysis, the optimal data transformation resulting from the BoxCox test (Box & Cox, 1964) was to raise RTs to the power of −.38. The initial linear mixed effect model included by-item and by-participant random intercepts, and trial order, preceding RT, visual complexity, imageability, word frequency, familiarity, length in phonemes, Australian speeded name agreement, number of years of exposure to English, phonological edit distance, and bilingual-specific English PND as fixed factors, as well as the interaction between phonological edit distance and exposure to English, the interaction between English exposure and English PND, and the interaction between target familiarity and bilingual-specific English PND (see Appendix B). Next, model comparison was performed to determine the significant contribution of each fixed factor and interaction in the model. Predictors that improved the model were name agreement ($X^2(1) = 85.65$, $p < .001$), response time of the preceding trial ($X^2(1) = 38.64$, $p < .001$), target familiarity ($X^2(1) = 35.25$, $p < .001$), and trial order ($X^2(1) = 20.46$, $p < .001$), and for the experimental predictors, phonological edit distance ($X^2(1) = 27.23$, $p < .001$). Predictors that did not significantly improve the model were log word frequency ($X^2(1) = 3.03$, $p = .082$), visual complexity ($X^2(1) = 2.71$, $p = .099$), length ($X^2(1) = 0.001$, $p = .974$)[4], and for the experimental predictors, exposure to English ($X^2(1) = 3.63$, $p = .057$), and bilingual-specific PND ($X^2(1) = 1.45$, $p = .228$).

The tested interactions all improved the model's fit: the interaction between phonological edit distance and the number of years of exposure to English ($X^2(1) = 5.47$, $p = .019$), between number of years of exposure to English and bilingual-specific English PND ($X^2(1) = 6.07$, $p = .014$), and between familiarity and bilingual-specific English PND ($X^2(1) = 10.16$, $p = .001$). Because imageability values were only available for 249 items, chi-square type model comparison was not possible, hence the independent contribution of imageability was assessed by considering the Akaike Information Criterion (AIC) of a model with imageability (−136,331.8) and a model without this variable (−171,489.7). A difference in AIC of more than two is substantial evidence of the superiority of the model with the lowest AIC (e.g. Burnham & Anderson, 2004), hence the model without imageability was deemed a better fit to the data than the model including imageability.

Next, we removed non-significant fixed factors (except for those that were included in significant interactions) which resulted in a model that included trial number, preceding RT, familiarity, Australian name agreement, English exposure, phonological edit distance and bilingual-specific English PND as fixed factors, and the three interactions. The resulting model is summarised in Table 4.

Consistent with previous studies of predictors of picture naming behaviour (e.g. Alario et al., 2004), in this final model, words named earlier in the experiment, with shorter preceding RTs, higher name agreement, and higher familiarity, were named faster (see Full model in Appendix C1). Interestingly, (English) log frequency was not a significant predictor of latencies, reinforcing the inadequacy of this measure for individuals with lower levels of English proficiency. In line with Sadat et al. (2014), words with higher phonological overlap with their translation equivalent were named faster. Exposure to English was only marginally significant, with participants with less exposure tending to be slower. There was a significant interaction between

Table 3. Summary of significant predictors and interactions for each response time and accuracy model corresponding to a given measure of PND, and model AIC.

Model corresponding to	Previous RT	Trial order	Imageability	Name agreement	Familiarity	Phonological edit distance	English exposure	PND measure	PND measure * Familiarity	PND measure * Exposure	Phonological edit distance * Exposure	Model AIC
Response time												
English PND (biling)	↗	↘	n/a	↗	↗	↘	(↗)	ø	✓	✓	✓	−171555.30
English PND (trad)	↗	↘	n/a	↗	↗	↘	(↗)	ø	✓	✓	✓	−171554.60
French PND (biling)	↗	↘	n/a	↗	↗	↘	(↗)	ø	✓	✓	✓	−171557.00
French PND (trad)	↗	↘	n/a	↗	↗	↘	(↗)	ø	✓	✓	✓	−171556.30
Combined PND (biling)	↗	↘	n/a	↗	↗	↘	(↗)	ø	✓	✓	✓	−171559.20
Combined PND (trad)	↗	↘	n/a	↗	↗	↘	(↗)	ø	✓	✓	✓	−171560.10
Accuracy												
English PND (biling)	n/a	↘	↘	n/a	↗	↘	↗	ø	✓	ø	ø	10890.77
English PND (trad)	n/a	↘	↘	n/a	↗	↘	↗	ø	✓	ø	ø	10891.07
French PND (biling)	n/a	↘	↘	n/a	↗	↘	↗	ø	✓	ø	ø	10891.49
French PND (trad)	n/a	↘	↘	n/a	↗	↘	↗	ø	✓	ø	ø	10894.63
Combined PND (biling)	n/a	↘	↘	n/a	↗	↘	↗	ø	✓	ø	ø	10887.87
Combined PND (trad)	n/a	↘	↘	n/a	↗	↘	↗	ø	✓	ø	ø	10887.60

(biling) = Bilingual-specific PND measures; (trad) = Traditionally computed PND measures; ↗ = facilitatory effect; ↘ = inhibitory effect; ø = no significant effect; ✓ = presence of a significant interaction.

Table 4. (Transformed) Response time analyses: Summary of the final Bilingual-specific PND: model structure, Akaike Information Criterion (AIC), Variance and standard deviation (SD) of random effects, coefficients (B) and their standard errors (SE), *t*-values, *p*-values, and variance inflation factor (VIF), for each fixed effect and interaction.

Model structure	Lmer (RTT ~ (1 | participant) + (1 | item) + standardised preceding RT + trial order + name agreement + familiarity + phonological edit distance * exposure to English + Bilingual-specific English PND * exposure to English + Bilingual-specific English PND * familiarity, data: French late bilinguals, optimiser: nloptwrap)				
AIC	−171555.3				
Random effects	Variance	SD			
Item (intercept)	1.17E-08	1.08E-04			
Participant (intercept)	1.12E-08	1.06E-04			
Residual	4.12E-08	2.03E-04			
Fixed effects	B	SE	*t*-value	*p*-value	VIF
(Intercept)	9.57E-04	1.63E-05	58.59	< .001	
Preceding RT standardised	−1.21E-05	1.95E-06	−6.22	<.001	1.00
Trial order	−1.11E-05	1.93E-06	−5.75	<.001	1.00
Name agreement	6.35E-05	6.48E-06	9.79	<.001	1.03
Familiarity	5.24E-05	6.56E-06	7.99	<.001	1.07
Bilingual-specific English PND	−6.63E-06	6.56E-06	−1.01	.313	1.06
English exposure	2.90E-05	1.51E-05	1.92	.061	1.00
Phonological edit distance	−3.55E-05	6.54E-06	−5.42	<.001	1.05
Familiarity * Bilingual-specific English PND	2.03E-05	6.60E-06	3.07	.002	1.04
English exposure * Bilingual-specific English PND	4.55E-06	1.85E-06	2.46	.014	1.01
English exposure * Phonological edit distance	4.51E-06	1.93E-06	2.34	.020	1.01

All variables in the fixed effects were standardised.

English exposure and phonological edit distance, as illustrated in Figure 1, Panel A: Participants were generally faster to name targets with more phonological overlap (low phonological edit distance) with their French translation equivalents than targets with less phonological overlap, a "cognateness benefit". However, the size of this benefit was larger for those participants with little English exposure than for participants with more English exposure. In other words, the "cognateness benefit" reduced as English exposure increased.

The measure of bilingual-specific English PND interacted with English exposure (see Figure 1, Panel B) such that individuals with more exposure to English exhibited a negligible effect of bilingual-specific PND while individuals with less exposure to English were slower with words of higher bilingual-specific PND than words with lower bilingual-specific PND.

Bilingual-specific English PND also interacted with familiarity (see Figure 1, Panel C): There was little effect of bilingual-specific English PND on higher familiarity items (if anything, the effect trended in the direction of facilitation), and the largest (inhibitory) effect of bilingual-specific English PND was seen on lower familiarity targets (lower familiarity words were slower named if they had higher bilingual-specific English PND).

Response time analyses – other PND measures

Subsequently, bilingual-specific English PND was replaced by every other phonological neighbourhood density measure (English PND tradit., bilingual-specific French PND, French PND tradit., bilingual-specific Combined PND, and Combined PND tradit.) in separate models. Effects of the different measures of PND, the other experimental variables and the different tested interactions, along with the AIC of each model, are summarised in Table 3, and full response time models are displayed in Appendix C.

Effects of the other PND measures were very comparable to the English PND measure: all the interactions discussed above for English PND were also significant for French PND and Combined (English & French) PND and the parallel measures using traditional calculations (rather than calculations adjusted for late bilingual phonology). As one of our main goals was to determine which PND measure was the best predictor of picture naming performance, we considered each model's AIC (models with at least a 2 units lower AIC are a better fit than models with a higher AIC: Burnham & Anderson, 2004). Since these models only differ by one predictor (the specific measure of PND), differences in AIC reflect the particular contribution of each of these critical predictors to the fit of the model. The models including the Combined PND measures provided the best fit of the data, with lower AICs than the models including the French PND measures, which in turn had lower AICs than English PND measures. Finally, contrary to our expectations, models featuring PND measures computed using our bilingual-specific measure were not meaningfully different from the corresponding models featuring traditionally computed PND measures.

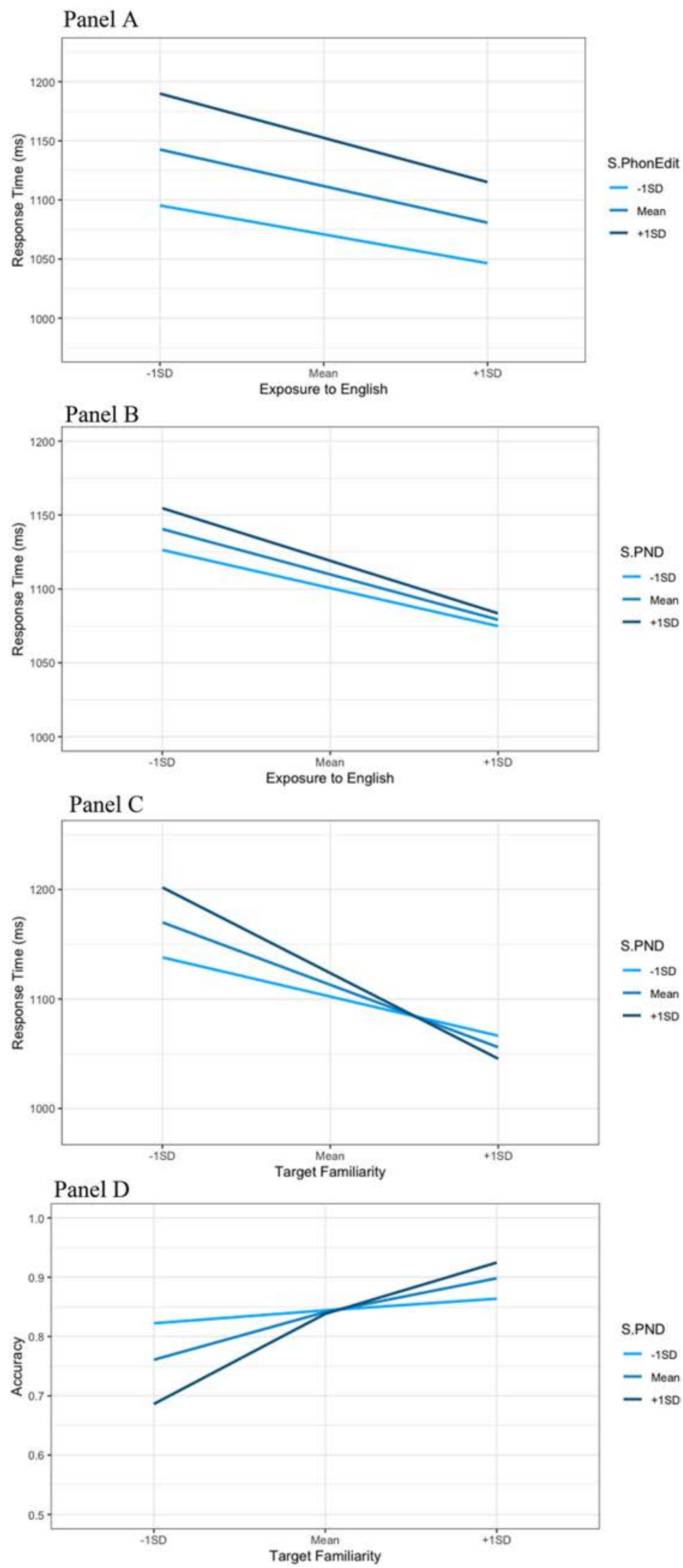

Figure 1. Illustration of significant interactions (all from Bilingual-specific English PND models). Panel A: Exposure to English * standardised phonological edit distance (S.Phon.edit) on response time. Panel B: Exposure to English * standardised Bilingual-specific English PND (S.PND) on response time. Panel C: Familiarity * standardised Bilingual-specific English PND (S.PND) on response time. Panel D: Familiarity * standardised Bilingual-specific English PND (S.PND) on accuracy.

Note: for plotting clarity, raw RTs were used in the response time models plots.

Accuracy analyses – (bilingual-specific) English PND

Overall accuracy was 77%. Accuracy (error = 0, correct = 1) was analysed using generalised linear mixed effect models, using the same procedure as for latency but excluding preceding RT, and name agreement (as name agreement here is simply a measure of mean accuracy per-item). Predictors that improved the model were familiarity ($X^2(1) = 26.47$, $p < .001$), trial order ($X^2(1) = 11.30$, $p < .001$), log frequency ($X^2(1) = 6.79$, $p = .009$) and for the experimental predictors, phonological edit distance ($X^2(1) = 7.30$, $p = .007$) and number of years of exposure to English ($X^2(1) = 13.83$, $p < .001$). The model with imageability had a lower AIC than the model without imageability (10,892.36 versus 14,165.56). Predictors that did not improve the model (as main or simple effects) were length ($X^2(1) = 0.024$, $p = .877$), visual complexity ($X^2(1) = 0.032$, $p = .858$), and for the experimental predictors, bilingual-specific English PND ($X^2(1) = 0.5357$, $p = .464$). The addition of the interaction between phonological edit distance and exposure to English did not improve the accuracy model ($X^2(1) = 0.88$, $p = .349$), and neither did the interaction between bilingual-specific English PND and exposure to English ($X^2(1) = 0.30$, $p = .586$). The interaction between familiarity and bilingual-specific English PND did improve the model ($X^2(1) = 16.57$, $p < .001$). The predictors retained in the model in addition to bilingual-specific English PND, were, therefore, trial order, familiarity, log frequency, imageability, English exposure and phonological edit distance, as well as the interaction between familiarity and bilingual-specific English PND. In this final model (summarised in Table 5), there were significant effects such that targets that were named earlier in the experiment, had higher imageability, higher familiarity ratings and higher frequency were named more accurately. In addition, targets that were lower in phonological edit distance (more similar phonological form) to their French translation equivalent were more accurately named and participants with more extensive exposure to English showed more accurate naming. Bilingual-specific English PND interacted with familiarity in the following way: for lower familiarity items, there was an inhibitory effect of bilingual-specific English PND, whereas on more familiar targets, there was no (or if anything, a slightly facilitatory) influence of bilingual-specific English PND[5] (see Figure 1, Panel D).

Accuracy analyses – other PND measures

Similar to response time analyses, other models were computed using the same structure as the final bilingual-specific English PND model but replacing this measure by the other PND measures, one at a time. All model summaries are available in Appendix D. Again, results were comparable across all PND measures (See Table 3 for a summary). We inspected AIC values for each accuracy model, and found that the models with the best fit to the data were those using combined measures of PND. The model that had the highest AIC (hence with the least good fit) was the one that included French PND computed the traditional way. Measures computed with our bilingual-specific approach were, once again, not meaningfully better than PND measures computed the traditional way (apart from French PND where the fit of the corresponding bilingual-specific model was better than that of the model that included the same measure but computed the traditional way).

Discussion

This experiment examined picture naming, in English, by French native speakers who had acquired English later in life but had been immersed in an English-speaking environment for at least the previous two years. We used a continuous measure of "cognateness", and a novel measure of phonological neighbourhood density in an attempt to capture the specific phonological representations of non-native words in a late bilingual's lexicon, and took into account the more limited range of vocabulary expected in these late bilinguals. We aimed to investigate to what extent phonologically related words in both languages of a late bilingual influenced the picture naming performance in L2, and whether familiarity of the target and number of years of exposure of the speaker to their second language influenced potential effects of phonological neighbourhood density.

The role of language exposure

Unsurprisingly, picture naming became more accurate and marginally faster the more extensive the participants' exposure to English (operationalised as years of immersion in an English speaking environment). We found the standard "cognate" facilitation effect on response time and accuracy but using a more nuanced (continuous) measure of phonological overlap (see also Sadat et al., 2016): the more similar sounding an English word was to its French counterpart, the faster and the more accurately it was named. In addition, we showed that this cognate facilitation effect on response time in L2 decreased with English exposure. This is in line with Costa et al. (2000) who found stronger cognate

Table 5. Accuracy analyses: Summary of the final Bilingual-specific PND: model structure, Akaike Information Criterion (AIC), Variance and standard deviation (SD) of random effects, coefficients (B) and their standard errors (SE), z-values, *p*-values, and variance inflation factor (VIF), for each fixed effect and interaction.

Model structure	glmer (accuracy ~ (1 | participant) + (1 | item) + trial order + imageability + familiarity + phonological edit distance + exposure to English + Bilingual-specific English PND * familiarity, data: French late bilinguals, optimiser: bobyqa)				
AIC	10890.77				
Random effects	Variance	SD			
item (intercept)	1.292	1.137			
participant (intercept)	0.277	0.526			
Fixed effects	B	SE	*z*-value	*p*-value	VIF
(Intercept)	1.67E+00	1.10E-01	15.22	<.001	
Trial order	−8.74E-02	2.56E-02	−3.41	.001	1.00
Imageability	3.38E-01	7.81E-02	4.33	<.001	1.03
Familiarity	5.10E-01	8.49E-02	6.01	<.001	1.07
English exposure	3.15E-01	7.92E-02	3.98	<.001	1.00
Bilingual-specific English PND	−2.19E-02	7.82E-02	−0.28	.779	1.05
Phonological edit distance	−2.49E-01	8.17E-02	−3.05	.002	1.07
Familiarity * Bilingual-specific English PND	3.53E-01	8.35E-02	4.23	<.001	1.09

All variables in the fixed effects were standardised.

effects when bilinguals name pictures in their L2 (less proficient language) compared to their L1 (more proficient language). However, our response time results extend this by demonstrating that these proficiency differences can be seen even within the non-dominant language, as proficiency increases with exposure.

Turning to the influence of language exposure on the effects of phonological neighbourhood, our results differ from those of Marian and Blumenfeld (2006): we found that measures of phonological neighbourhood density had an inhibitory effect on word production in the individuals with the shortest exposure to English, while Marian and Blumenfeld found stronger facilitatory effects of phonological neighbourhood density when the target language was the second, less proficient language (as opposed to the first language). However, our participant sample likely had lower proficiency overall than that of Marian and Blumenfeld (2006). Marian and Blumenfeld's participants were proficient late bilinguals, with longer exposure to their second language compared to our study (these participants had started "learning" their second language 17 years on average prior to the experiment, while our sample had been exposed to the second language for as little as two years, and for about ten years on average). Hence, Marian and Blumenfeld's sample is more similar to those individuals in our study who had the longest exposure. The fact that, in our study, an interaction was found between the effect of (within language) phonological neighbourhood density and language exposure fits with this interpretation: in the present study the inhibitory effect of neighbours was only present for participants with less exposure to English, but it evolved towards a facilitatory trend for those participants whose exposure time was possibly more similar to that of Marian and Blumenfeld's participants.

However our finding of a significant interaction between exposure and phonological neighbourhood density measures does not fit with Sadat et al. (2016) who found main inhibitory effects of target language phonological neighbourhood density and no effect of cross-language phonological neighbourhood density in highly proficient early bilinguals. Indeed, our findings suggested that highly proficient bilinguals should show a trend towards *facilitatory* effects of (within language) phonological neighbourhood density. We discuss the potential source of this difference below.

Effects of PND depend on target familiarity

Aside from effects of L2 proficiency on the direction of phonological neighbourhood density effects on latency, we also found interactions between every measure of phonological neighbourhood and familiarity, in both response time and accuracy: There was an inhibitory effect of higher phonological neighbourhood density but only for more unfamiliar words. In contrast, for words of higher familiarity, the effect, if anything, tended to be facilitatory. If this is the case, then depending on the overall familiarity of an item set, we might see different effects of phonological neighbourhood density emerging in L2 picture naming, that is, inhibitory effects when the dataset is of overall low familiarity and facilitatory effects if the dataset is of particularly high familiarity. This possibility was examined by retrieving concept familiarity values corresponding to Marian and Blumenfeld (2006) and Sadat et al.'s (2016) stimuli. We took the English translations of items from both studies (Marian and Blumenfeld: German; Sadat et al.: Spanish) and retrieved the corresponding familiarity values using the same source as in the present study (Johnston et al., 2010). Sixty-eight percent of the English counterparts of Marian and Blumenfeld's items were in Johnston et al.'s database, and 71% of Sadat et al.'s items. Familiarity values were significantly higher in Marian et al.'s dataset (mean = 5.48, SD = 1.04) compared to ours (mean = 5.02, SD = 1.15); ($t(355) = -2.37$, $p = .018$, two-tailed) and to Sadat et al.'s (mean = 4.90, SD = 1.16); ($t(418) = -2.98$, $p = .001$, two-tailed). Familiarity values in our dataset did not differ significantly from those in Sadat et al. ($t(697) = 1.35$, $p = .176$, two-tailed). These results are consistent with the hypothesis that, in bilingual picture naming, higher familiarity datasets can generate facilitatory PND effects (as is the case in Marian & Blumenfeld, 2006) and that low familiarity datasets are more likely to generate inhibitory PND effects (as in Sadat et al., 2016). Of course, we must be cautious when interpreting these results as only around 70% of the familiarity values were available: these values might not be fully representative of the complete sets.

Late bilinguals are affected by both within – and cross-language phonological neighbourhood density

Effects were observed in the same direction for all three measures of phonological neighbourhood density (within- and cross- language and combined). This suggests that at least in late bilinguals with limited proficiency, phonological neighbours (of both languages) are activated during word production and influence processing. This section addresses the fact that the measures relating to Combined and French (cross-language, L1) neighbourhood appeared to be slightly better predictors than those pertaining to English (within-language, L2) neighbours. We will discuss this, while being aware that some caution needs to be used: the three mesures were intercorrelated, and it is therefore difficult to be confident what the contribution of each measure to the observed effects was.

The relative apparent superiority of Combined PND and French (cross-language) PND, together with the fact that, at least on response time, effects of phonological neighbourhood density were stronger in individuals with lowest levels of exposure, suggests that the effects may be driven by neighbours of the most proficient or dominant language (here, in the least balanced individuals, French). As dominance changes, English words may have a stronger impact on lexical processing, and the opposite occurs for French words. That is, the relative familiarity difference between English targets and French neighbours decreases as proficiency increases and dominance changes. Marian and Blumenfeld (2006) argued that stronger effects of phonological neighbourhood density on response time in lower proficiency individuals could be attributed to the fact that words in a less proficient language have an overall "low frequency status" that makes these words "more sensitive" to phonological neighbourhood density (of that language). We agree that words of a less proficient language will have an overall lower frequency status, but as this is the case for both the targets and their neighbours in that language, it is unclear why the relative balance of activation between target and neighbours would be different in low proficiency compared to high proficiency speakers. We argue that, as Marian and Blumenfeld proposed, in (late) bilinguals, words in a less proficient language are more sensitive to effects of phonological neighbours, but our view is that this is the (relatively higher "frequency") neighbours of the more proficient/dominant language that drive this effect.

An important point we want to acknowledge is that, following the language mode theory (e.g. Grosjean, 1998, 2001) the setting of our experiment may have promoted a higher activation of the non-target language. Indeed, participants were aware that the experimenter was proficient in French (as well as English). This likely promoted a higher activation of the non-target language (French), and hence might have contributed to enhanced cross-linguistic effects (such as effects of cross-language neighbours and cognates) compared to a strictly monolingual setting. There is therefore the possibility that the patterns we are seeing here might correspond to an "enhanced" version of cross-language dynamics in late bilinguals' lexical processing.

We note that analyses performed using PND measures computed without the "bilingual-specific" adaptation showed the same significant interactions on response time and accuracy (and no additional effects) compared to those performed with the adaptation. Even though, in general, models with these traditionally computed measures had a very slightly worse fit (higher AICs) than models with the bilingual-specific adaptation, these differences were too small to claim that the bilingual-specific measures provided a better fit to the data .While this finding was not expected, it shows the robustness of the observed PND effects in the present study. It also alleviates our concern, and that of Marian and Blumenfeld (2006) and Vitevitch (2012), that traditionally computed phonological neighbourhood measures may not be meaningful for bilinguals. Hence, these "standard" measures of phonological neighbourhood density might be just as effective at predicting bilingual picture naming performance as measures adapted to specific bilinguals.

How could the effects of phonological neighbourhood density in late bilinguals be interpreted within theories of speech production? In the case of monolinguals, it has been argued (e.g. Chen & Mirman, 2012; Dell & Gordon, 2003) that an influence of phonological neighbourhood density can be best explained by models that postulate cascading and interactivity between the lexical level and the phonological level: A facilitatory effect of neighbours arises from feedback from the target's phonemes to both the target and its phonological neighbour's lexical representation. This in turn leads to greater activation of the target phonemes and stronger activation of the target's lexical node resulting in higher likelihood of selection. Hence, the more neighbours, the more likely the target is to be selected and the more extra activation is sent to target phonemes. Note that, although both Chen and Mirman (2012) and Dell and Gordon (2003) only envisaged facilitatory effects of phonological neighbours, inhibitory effects could be possible within the same account: If phonological neighbours are strongly activated at the lexical level, approaching the level of activation of the target, then these neighbours might generate inhibitory effects. For instance, more time steps may be required for the target to reach a level of activation sufficiently superior to that of the phonological neighbour(s) to be selected, resulting in slower "naming" times for words with many phonological neighbours. Hence, in order to accommodate our results, a theory requires a mechanism that will result in both facilitatory and inhibitory effects of neighbours. The influence of phonological neighbourhood density in bilinguals (especially cross-language) has not received as much attention to date. There is strong evidence that, in bilingual speech production, activation is language-independent, that is, the non-target language is activated, at least to the level of phonological representations via the translation equivalent of the target word (Colomé, 2001; Rodriguez-Fornells et al., 2005). For instance, when naming a picture in one language, activation from semantics flows to the lexical nodes of both the target language and the non-target language, then each of these lexical nodes send activation to their respective phoneme nodes. However, it is one step further to assume activation of phonological neighbours of the other language, because it implies that phoneme nodes from the non-target language can also be activated via a lexical representation from the target language, and not only via activation of its translation equivalent. This entails that, in bilinguals, activation from a lexical representation that is specific to one language is sent to phoneme nodes in a language non-specific way (which means that the phonological repertoire is shared between languages). Furthermore, it implies that these phoneme nodes send activation back to lexical representations of *both* languages, meaning there is interactivity between levels of processing, both within and across languages (at least in the case of L2 naming). Based on the observation of larger cognate facilitation when the response was in the non-dominant language, Costa et al. (2000) argued that the amount of activation of the phoneme nodes level was proportional to the lexical node's activation level, which was higher for representations from a dominant language, as they are of higher frequency. Following this logic, it is then possible that for participants with less exposure to English (hence most probably dominant in French), when naming in English, French representations are relatively more highly activated, compared to participants with more extensive experience in English. Therefore, in our experiment, French neighbours could be serious competitors of an English target, resulting in an inhibitory effect of these French neighbours and of the Combined measure including both French and English neighbours. But when language dominance is shifted to the target language, less effect of (cross-language) neighbours might be expected, as we found in this experiment.

Costa and colleagues' argument of stronger activation of words in a dominant language also fits well with our finding that familiarity levels of the target (that evolve as proficiency increases) result in different patterns of phonological neighbourhood density effects.

Conclusion

This research investigated the patterns of (cross-) language activation in late bilinguals with a focus on the phonological properties of the bilingual lexicon. It showed that the influence of phonological neighbourhood density in English picture naming for late bilingual speakers is modulated by several factors relating to the relative frequency status of the target compared to that of its neighbours. When the target was low in familiarity the effect of phonological neighbours tended to be inhibitory, and facilitatory when the target word had high familiarity. In addition, phonological neighbourhood density measures tended to have an inhibitory effect on latencies in individuals with the lowest number of years of exposure to English compared to individuals with more exposure. A parallel can be drawn between these results and the predictions of the weaker links hypothesis (Gollan et al., 2008) in that language exposure could be equated to frequency status of words of the languages of a bilingual. When individuals had low exposure, target items were overall low in frequency/familiarity, with resultant inhibitory effects of phonological neighbourhood density, whereas individuals with higher exposure showed a trend for facilitatory effects of phonological neighbourhood density. We suggested that increased phonological neighbourhood density effects in individuals of lower exposure were driven by cross-language neighbours from the more proficient, dominant language. Moreover, we found that effects of form overlap between translation equivalents ("cognateness") were stronger in individuals with less exposure, showing again the influence of L2 exposure on patterns of cross-language activation. Taken together, these results suggest that the weaker links hypothesis can not only explain effects at the lexical level, but also effects that happen as a consequence of interactivity between the lexical and post-lexical/phoneme level of processing, within and across languages. We suggest that our results can be best accommodated within interactive activation models with interactivity between the lexical level and the phoneme level, and that in bilingual speakers activation and interactivity exist within and across languages. Finally, this study illustrates the need to incorporate effects of increasing exposure to L2 in the patterns of activation of the two languages within (late) bilingual speakers in bilingual models of spoken word production, and underlines the dynamic nature of the bilingual lexicon.

Notes

1. For the sake of simplicity, we refer to the population under investigation as late bilinguals or just bilinguals (other ways to refer to our population would be second language -or L2- learners). We are acutely aware of the debate around what constitutes a bilingual, and variability within this population and we do not intend to imply that our findings will necessarily generalise to all (late) bilingual speakers.
2. Forty English monolingual students from Macquarie University (Hameau, 2016; Hameau et al., 2018).
3. In CLEARPOND, French neighbours are drawn from the Lexique 2 database (New et al., 2004), using a frequency cut-off that ensures that the French neighbours are likely to be words from an educated French monolingual adult's vocabulary (i.e. excluding rare words that potentially very few participants would know).
4. Note that in a previous version of this manuscript (as part of the first author's PhD thesis: Hameau, 2016), we reported analyses run both with and without word length in phonemes, despite the absence of an effect on either response time or accuracy of this variable. The presence or absence of length as a predictor did not change the pattern of significance (interactions) of the experimental predictors.
5. Since Log frequency significantly predicted accuracy (as opposed to response time), we tested the interaction between frequency and PND in line with Hameau et al. (2018): this interaction did not improve the PND model: $X^2(1) = 0.0176$, $p = .894$.

Acknowledgements

We thank Dr Anthony Shook for helpful communication regarding the CLEARPOND database. During the preparation of this paper, the first author was funded by an Australian Research Council Discovery Project grant (190101490), and by an International Macquarie University Research Excellence Scholarship (iMQRES) in a previous version of this manuscript. In addition, we thank two anonymous reviewers for their useful suggestions aimed at improving the manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Australian Research Council [grant number DP 190101490] and an International Macquarie University Research Excellence Fellowship (iMQRES) to Solène Hameau.

ORCID

Solène Hameau http://orcid.org/0000-0001-5282-2645
Britta Biedermann http://orcid.org/0000-0001-6242-1167
Lyndsey Nickels http://orcid.org/0000-0002-0311-3524

References

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*(1), 140–155. https://doi.org/10.3758/BF03195559

Allison, P. (2012). *When can you safely ignore multicollinearity? | Statistical horizons*. Statistical Horizons. https://statisticalhorizons.com/multicollinearity

Baayen, H. R., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*(2), 12–28. http://www.redalyc.org/html/2990/299023506003, https://doi.org/10.21500/20112084.807

Baayen, R. H., Piepenbrock, R., & Rijn, v. H. (1993). *The {CELEX} lexical data base on {CD-ROM}*.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*. http://arxiv.org/abs/1406.5823

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and* *linguistic experience: Issues in cross-language research* (pp. 171–204). York Press.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O.-S. Bohn, & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Fledge* (pp. 13–34). John Benjamins Publishing Company.

Bonin, P., Chalard, M., Méot, A., & Fayol, M. (2002). The determinants of spoken and written picture naming latencies. *British Journal of Psychology*, *93*(1), 89–114. https://doi.org/10.1348/000712602162463

Box, G. E. P., & Cox, D. R. (1964). An analysis of Transformations. In *Journal of the royal statistical society. series B (methodological)* (Vol. 26, pp. 211–252). WileyRoyal Statistical Society. https://doi.org/10.2307/2984418

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods and Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Chen, X., & Dong, Y. (2019). Evaluating objective and subjective frequency measures in L2 lexical processing. *Lingua. International Review of General Linguistics. Revue internationale De Linguistique Generale*, *230*, 102738. https://doi.org/10.1016/J.LINGUA.2019.102738

Chen, Q., & Mirman, D. (2012). Competition and cooperation among similar representations: Toward a unified account of facilitative and inhibitory effects of lexical neighbors. *Psychological Review*, *119*(2), 417–430. https://doi.org/10.1037/a0027175

Colomé, À. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language*, *45*(4), 721–736. https://doi.org/10.1006/jmla.2001.2793

Colomé, À, & Miozzo, M. (2010). Which words are activated during bilingual word production? *Journal of Experimental Psychology: Learning Memory and Cognition*, *36*(1), 96–109. https://doi.org/10.1037/a0017677

Coltheart, M., Jonasson, J. T., Davelaar, E., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and performance VI*. Academic P. (pp. 64-85). Erlbaum.

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

Costa, A., Roelstraete, B., & Hartsuiker, R. J. (2006). The lexical bias effect in bilingual speech production: Evidence for feedback between lexical and sublexical levels across languages. *Psychonomic Bulletin & Review*, *13*(6), 972–977. https://doi.org/10.3758/BF03213911

Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language*, *94*(1), 94–103. https://doi.org/10.1016/j.bandl.2004.12.002

de Groot, A. M. B., Borgwaldt, S., Bos, M., & van den Eijnden, E. (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language*, *47*(1), 91–124. https://doi.org/10.1006/JMLA.2001.2840

Dell, G. S., & Gordon, J. K. (2003). Neighbors in the lexicon! Friends or foes? In N. O. Schiller, & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production* (pp. 9–47). Mouton de Gruyter.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, *104*(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Fledge, J. E. (1995). Second language speech learning. Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.

Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, *35*(1), 116–124. https://doi.org/10.3758/BF03195503

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression*. https://cran.r-project.org/web/packages/effects/effects.pdf

García Lecumberri, M. L., & Cenoz Iragui, J. (1997). L2 perception of English vowels: Testing the validity of Kuhl's prototypes. *Revista Alicantina de Estudios Ingleses*, *10*(10), 55–68. https://doi.org/10.14198/raei.1997.10.06

Garlock, V. M., Walley, A. C., & Metsala, J. L. (2001). Age-of-acquisition, word frequency, and neighborhood density effects on spoken word recognition by children and adults. *Journal of Memory and Language*, *45*(3), 468–492. https://doi.org/10.1006/JMLA.2000.2784

Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, *113*(2), 256–281. https://doi.org/10.1037/0096-3445.113.2.256

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, *58*(3), 787–814. https://doi.org/10.1016/j.jml.2007.07.001

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, *1*(2), 131–149. https://doi.org/10.1017/S136672899800025X

Grosjean, F. (2001). The bilingual's language modes. In J. L. Nicol (Ed.), *One mind, two languages: Bilingual language processing* (pp. 1–22). Blackwell. https://doi.org/10.1002/9781405198431

Hall, K. C., Allen, B., Fry, M., Mackie, S., & McAuliffe, M. (2015). *Phonological corpus tools, version 1.1.0*. https://github.com/PhonologicalCorpusTools/CorpusTools/releases

Hameau, S. (2016). *Neighbourhood density effects in spoken word production*. Macquarie University.

Hameau, S., Biedermann, B., & Nickels, L. (2018). Effects of phonological neighbourhood density and frequency in picture naming: Australian and British English, and computational modelling. *10th International Workshop on Language Production*.

Hutcheson, G., & Sofroniou, N. (1999). *The multivariate social scientist: introductory statistics using generalized linear models*. Sage Publications.

Iverson, P., & Evans, B. G. (2007). Learning English vowels with different first-language vowel systems: Perception of formant targets, formant movement, and duration. *The Journal of the Acoustical Society of America*, *122*(5), 2842–2854. https://doi.org/10.1121/1.2783198

Johnston, R. A., Dent, K., Humphreys, G. W., & Barry, C. (2010). British-English norms and naming times for a set of 539 pictures: The role of age of acquisition. *Behavior Research Methods*, *42*(2), 461–469. https://doi.org/10.3758/BRM.42.2.461

Kirsner, K., Lalor, E., & Hird, K. (1993). The bilingual lexicon. In R. Schreuder, & B. Weltens (Eds.), *Studies in bilingualism, Vol. 6. The bilingual lexicon* (pp. 215–248). John Benjamins Publishing Company. https://doi.org/10.1075/sibil.6.11kir

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). Lmertest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), https://doi.org/10.18637/jss.v082.i13

Luce, P. A. (1987). Neighborhoods of words in the Mental lexicon. *Dissertation Abstracts International, B: Sciences and Engineering*, *47*(12), https://eric.ed.gov/?id=ED353610

Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*(1), 1–36. http://www.ncbi.nlm.nih.gov/pubmed/9504270 https://doi.org/10.1097/00003446-199802000-00001

Luce, P. A., Pisoni, D. B., & Goldinger, S. D. (1990). Similarity neighborhoods of spoken words. In G. T. M. Altmann (Ed.), *ACL–MIT Press series in natural language processing. Cognitive models of speech processing: Psycholinguistic and computational perspectives* (pp. 122–147). The MIT Pr. http://psycnet.apa.org/record/1991-97555-005

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

Marian, V., & Blumenfeld, H. K. (2006). Phonological neighborhood density guides: Lexical access in native and non-native language production. *Journal of Social and Ecological Boundaries*, *2.1*(1), 3–35.

Morton, J. (1970). A functional model for memory. In D. A. Norman (Ed.), *Models of human memory* (pp. 203–254). Academic Press.

Mulatti, C., Reynolds, M. G., & Besner, D. (2006). Neighborhood effects in reading aloud: New findings and new challenges for computational models. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(4), 799–810. https://doi.org/0.1037/0096-1523.32.4.799

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516–524. https://doi.org/10.3758/BF03195598

Nickels, L., & Howard, D. (1995). Aphasic naming: What matters? *Neuropsychologia*, *33*(10), 1281–1303. https://doi.org/10.1016/0028-3932(95)00102-9

Pallier, C., Colomé, À, & Sebastián-Gallés, N. (2001). The influence of native-language phonology on lexical access: Exemplar-based versus abstract lexical entries. *Psychological Science*, *12*(6), 445–449. https://doi.org/10.1111/1467-9280.00383

Protopapas, A. (2007). Check vocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods*, *39*(4), 859–862. https://doi.org/10.3758/BF03192979

R Development Core Team. (2011). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rodriguez-Fornells, A., Van Der Lugt, A., Rotte, M., Britti, B., Heinze, H. J., & Münte, T. F. (2005). Second language interferes with word production in fluent bilinguals: Brain potential and functional imaging evidence. *Journal of Cognitive Neuroscience*, *17*(3), 422–433. https://doi.org/10.1162/0898929053279559

Sadat, J., Martin, C. D., Costa, A., & Alario, F.-X. (2014). Reconciling phonological neighborhood effects in speech production through single trial analysis. *Cognitive Psychology*, *68*, 33–58. https://doi.org/10.1016/J.COGPSYCH.2013.10.001

Sadat, J., Martin, C. D., Magnuson, J. S., Alario, F.-X., & Costa, A. (2016). Breaking down the bilingual cost in speech production. *Cognitive Science*, *40*(8), 1911–1940. https://doi.org/10.1111/cogs.12315

Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, *1*(2), 103–113. https://doi.org/10.1111/j.2041-210X.2010.00012.x

Starreveld, P. A., De Groot, A. M. B., Rossmark, B. M. M., & Van Hell, J. G. (2014). Parallel language activation during word processing in bilinguals: Evidence from word production in sentence context. *Bilingualism*, *17*(2), 258–276. https://doi.org/10.1017/S1366728913000308

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

Székely, A., & Bates, E. (2000). Objective visual complexity as a variable in studies of picture naming. *Center for Research in Language Newsletter*, *12*(2), 3–33.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Lu, C. C., Pechmann, T., Pléh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., … Bates, E. (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, *51*(2), 247–250. https://doi.org/10.1016/j.jml.2004.03.002

Test your vocab. (2013). http://www.testyourvocab.com

Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary development. *International Journal of Bilingualism*, *15*(4), 426–445. https://doi.org/10.1177/1367006911403202

van Hell, J. G., & de Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, *1*(3), 193–211. https://doi.org/10.1017/s1366728998000352

van Heuven, W. J. B., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language*, *39*(3), 458–483. https://doi.org/10.1006/JMLA.1998.2584

Vitevitch, M. S. (2012). What do foreign neighbors say about the mental lexicon? *Bilingualism: Language and Cognition*, *15*(1), 167–172. https://doi.org/10.1017/S1366728911000149

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in perception of spoken words. *Psychological Science*, *9*(4), 325–329. https://doi.org/10.1111/1467-9280.00064

Vitevitch, M. S., & Luce, P. A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language*, *40*(3), 374–408. https://doi.org/10.1006/JMLA.1998.2618

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://cran.r-project.org/web/packages/ggplot2/ggplot2.pdf