Integration of newly learned L2 words into the mental lexicon is modulated by vocabulary learning method

Gary Boddaert, Camille Cornut, Séverine Casalis

Acta Psychologica 212 (2021) 103220

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Integration of newly learned L2 words into the mental lexicon is modulated by vocabulary learning method

G. Boddaert *, C. Cornut, S. Casalis

Univ. Lille, UMR 9193 - SCALab – Sciences Cognitives et Sciences Affectives, F-59000 Lille, France
CNRS, UMR 9193, F-59000 Lille, France

ARTICLE INFO

Keywords:
L2 learning
Prime lexicality effect (PLE)
Learning method
Novel word lexicalisation

ABSTRACT

The aim of the study was to investigate both L2 word integration and the effect of learning method on it. For this purpose, an L2 word-learning paradigm was designed with two learning methods: L2 words were paired with videos in the first one and their translation-equivalent L1 words in the second. To test L2 word integration, a lexical decision task associated with form priming was administered before and after the learning phase. The L2 words to be learned were used as primes. Forty-eight participants participated in the study. Before learning, a facilitation effect was obtained with pseudowords (not already learned L2 words) as primes and L1 words as targets. After learning, L2 words no longer facilitated L1 word recognition when learned with the video method, while they still had this effect when learned with the L1 words – L2 words method. In accordance with the prime lexicality effect (PLE), this absence of a facilitation effect indicates that L1 words and L2 words are involved in a lexical competition process common to the two languages. This result highlights swift lexicalisation and demonstrates the effect of learning method in lexicalisation.

1. Introduction

One of the features of second language (L2) learning in adulthood is that the first language (L1) is already clearly established when the L2 starts to be learned. Adults must learn new lexical forms (L2 words), while they have already learned lexical forms (L1 words) for the concepts depicted by the L2 words. Hence, important questions arise regarding how L2 words are added to memory in adults and then stored. One of these questions is how L2 words are integrated with L1 words at the beginning of L2 learning. This raises the issue of a separate vs integrated lexicon for words from different languages. Another issue is the role of learning method in L2 word integration.

In most bilingual models of visual word recognition, L1 and L2 words form part of a shared lexicon. For example, the Bilingual Interactive Activation plus model (BIA+, Dijkstra & Van Heuven, 2002) and the recent Multilink model (Dijkstra et al., 2018) posit a non-selective lexical access. In these models a visual presentation of a word leads to a co-activation and to a lexical competition between many word candidates that are similar to the input independently of the language they belong to. Interestingly, the fact that lexical representations are engaged in a lexical competition process can be used to investigate lexical integration of newly learned words since to be engaged in this process is a strong marker of lexical integration.

With this in mind the integration of newly learned words can be investigated through neighbourhood effect (e.g. Meade et al., 2018) or with the prime lexicality effect (Forster & Veres, 1998, see also Davis & Lupker, 2006). Under a masked priming condition, a target word (CONVERGE) is recognised faster when preceded by a form-related pseudoword sharing all but one letter (convenge CONVERGE) than an unrelated pseudoword (basoball CONVERGE). By contrast, a target word (CONVERGE) is recognised more slowly when preceded by a related word (converse CONVERGE) than an unrelated one (baseball CONVERGE). This so-called prime lexicality effect shows that words compete with each other. Indeed, given that pseudowords lead to facilitation and no inhibition, the competition effect is located at the lexical level.

Bilingual participants display a cross-language competition effect. Infrequent L1 words were recognised more slowly when preceded by a form-related frequent word than by an unrelated frequent word, whatever the language of the prime word: L1 (lire-CIRE vs nuit-CIRE) or L2 (fire-CIRE vs fall-CIRE) (Bijeljac-Babic et al., 1997). This effect was modulated by L2 proficiency, with highly proficient bilinguals

* Corresponding author at: Laboratoire SCALab UMR CNRS 9193, Université Lille 3, BP 60149, 59653 Villeneuve d'Ascq Cedex, France.
E-mail address: gary.boddaert@me.com (G. Boddaert).

https://doi.org/10.1016/j.actpsy.2020.103220
Received 5 March 2020; Received in revised form 22 July 2020; Accepted 11 November 2020
Available online 25 November 2020

displaying a stronger inhibition effect. Note that even though Multilink includes a proficiency component, this model (as well as BIA+) deals with word recognition in proficient bilinguals.

Besides, word learning has been described in developmental models of bilingual lexical organisation (see for example: Kroll & Stewart, 1994, see also Dong et al., 2005, SAM; Pavlenko, 2009, MHM). In these models L1 and L2 words are stored in separate lexicons and direct links connect L1 and L2 translation equivalents. A direct connection from L2 words to the conceptual level is thought to be established progressively as proficiency increases. The BIA-d (BIA-d, Grainger et al., 2010) provided a new interpretation of L2 word learning in which L1 and L2 words are first connected and belong to different lexicons. As proficiency increases, words from the two lexicons are progressively integrated and L2 word connections with semantics become stronger. Therefore, the former organisation first corresponds to RHM but becomes a BIA+ organisation as proficiency increases. Consequently, according to RHM no cross-language PLE is expected, while according to BIA-d a cross language PLE should be observed in more advanced L2 learners but not in novice learners (for an in-depth discussion about these models see Meade & Dijkstra, 2017). A few studies investigated how newly learned words are integrated in the lexicon.

According to Gaskell and Dumay (2003), integration of newly learned words does not occur immediately after learning but is observable the day after it, suggesting that lexicalisation depends on a consolidation process requiring exchanges between the medial temporal lobe system (including hippocampus) and the neocortical system (see also Davis & Gaskell, 2009). These exchanges are thought to take place during offline periods such as sleep. Therefore, sleep might play an important role in lexicalisation (for lexicalisation of words learned orally, see for example: Dumay & Gaskell, 2007; for lexicalisation of words learned through written modality, see for example: Wang et al., 2017 but see the following article for integration of spoken words without sleep: Lindsay & Gaskell, 2013). Critically, Qiao and Forster (2013) investigated new written word integration through the PLE. They found that newly learned L1 words used as primes did not display any L1 word detection facilitation effect after four training sessions spread over two weeks. The participants learned 48 new words (pseudowords) created by changing one letter in an English word. After the learning sessions, they completed a lexical decision task (LDT) under a masked priming condition with both learned items and unlearned items as primes. A significant facilitation effect was observed only with the unlearned items. Therefore, the newly learned L1 words had developed their own lexical representations. The question, therefore, is whether the process of integrating newly learned L2 words is similar to that of integrating L1 words.

Unfortunately, results concerning the integration of L2 words are mixed. Elgort (2011) and Elgort and Piasecki (2014) showed a PLE with L2 words, while Qiao and Forster (2017) did not. In Elgort's experiment (2011), 48 advanced non-native speakers of English (the participants' L1 was not controlled) learned 48 new English words (pseudowords created by changing one letter in an English word and presented as L2 words). After learning, these new L2 words (learned pseudowords) did not facilitate L2 word recognition, contrary to non-learned pseudowords (facilitation of 61 ms). Similar results were obtained in the study by Elgort and Piasecki (2014) where 48 Dutch-English bilinguals learned 48 new English words (pseudowords created by changing one letter in an English word and presented as L2 words). After learning, pseudowords learned as L2 words did not facilitate recognition of phonologically close already known L2 words, contrary to non-learned pseudowords (facilitation of 46 ms). Nevertheless, these results are challenged by those of Qiao and Forster (2017). They used the same procedure as in 2013 (Qiao & Forster) and found that newly learned L2 words (pseudowords learned as L2 words) facilitated the detection of already known L2 words (i.e. PLE not observed) in Chinese-English bilinguals. They interpreted this result as proof that "L2 words are stored in a different memory system from L1 words" (see also: Jiang & Forster, 2001). Thus, they hypothesized that L1 words are stored in the semantic memory while L2 words are stored in the episodic memory.

In summary, Elgort's (2011) and Elgort and Piasecki's experiments (2014) suggest that the process of integrating L2 words is similar to that of L1 words, while Qiao and Forster's experiment (2017) suggests a different integration process for L1 words and L2 words. This difference might be due to various factors, particularly differences in prime visibility, since presentation durations were different: 522 ms in Elgort (2011), 490 ms in Elgort and Piasecki (2014), and 50 ms in Qiao and Forster (2017). It might also be due to participants' characteristics and especially differences between participants' L1 and L2. In Qiao and Forster's experiment (2017), the two languages of the participants did not share a writing system: the L1 of the participants had a logographic writing system (Chinese), while the second one had an alphabetic one (English). On the other hand, in Elgort and Piasecki's experiment (2014), the two languages shared the same alphabet (L1 was not controlled in Elgort, 2011). In short, we need more information to provide a clear picture of this issue.

Considering new word learning, especially in a new language, an aspect that has been examined widely in educational studies but only little in cognitive psychology is the contribution of learning methods in establishing connections between words and concepts. Previous studies conducted in psychology (Comesaña et al., 2009;Comesaña et al., 2010 ; Comesaña et al., 2012) already investigated the effect of learning method on L2 words processing. The results obtained were mixed. Indeed, in 2009, the results obtained by Comesaña et al. suggested that the L2-picture method promotes conceptual links, while the results of the two other experiments suggest either that the interference effect was higher with the translation equivalent method (Comesaña et al., 2010) or that the effect was not significantly different between the two groups (Comesaña et al., 2012). As pointed out by Comesaña and colleagues (Comesaña et al., 2010; Comesaña et al., 2012) the inclusion of cognates in these latter experiments could have led the participants to use a strategy based on orthographic and phonological similarity between words independently of the learning method. Therefore, we need to know more about the way the L2 word learning method might impact whether or not words are integrated within the lexicon.

The main goal of the present experiment was to examine the integration process of L2 words into the mental lexicon in the first steps of learning, i.e. within the scope of learning new words of a new language, as a function of the method. The first objective was to investigate whether once integrated L2 representations can interact with L1 lexical representations. More specifically, the aim was to investigate whether lexical representations of newly learned L2 words are engaged in a lexical competition process shared by the two languages. The second objective was to investigate whether the learning method influences L2 words integration in the mental lexicon. For this purpose, we used a learning experiment in which 40 words of a new language were learned through two methods: L2 words were paired to videos in the first one, and their translation equivalent L1 words in the second one. The aim was to investigate the effect of a soft immersion in an L2 environment by presenting videos and L2 words without L1 words. This soft immersion is expected to lead to a deeper semantic processing and therefore to enhance lexical integration. To address the issue of lexical competition, these words were used as primes in an LDT with L1 (French) words as targets. Before the learning phase, a facilitation form priming effect is expected. But after learning the pattern expected is different in function of the hypothesis selected. If L2 words are not directly integrated into a single lexicon shared by the L1 and the L2 (i.e. hypothesis postulated for instance by RHM and BIA-d), the facilitation form priming would still be observed. But, if L2 words are integrated in the same lexicon as L1 words from the first steps of learning, a suppression of this facilitation effect is expected. Since the video method was expected to enhance lexical integration, the suppression of the facilitation effect is more likely to be observed in the video group.

2. Method

2.1. Participants

Forty-eight French native speakers studying in the University of Lille took part in the experiment (31 females, mean age = 23.7, SD = 3.5). The number of participants required was determined by taking into account the number of stimuli (40) and the number of participants used in previous comparing learning methods (respectively 48, 42 and 42 in Comesaña et al., 2009; Comesaña et al., 2010; Comesaña et al., 2012).

All participants have learned English at school at least from grade 6 and reached a medium level of proficiency in English (according to the scores obtained to the LEXTALE, Lemhöfer & Broersma, 2012, their average English level was B2. Mean English Lextale percentile and standard deviation for each group are reported in Table 1) and had received teaching in at least one other foreign language during schooling (mainly Spanish or German). Participants were randomly assigned to one of the two groups (one for each method). The two groups were homogeneous in French level (LEXTALE-FR, Brysbaert, 2013), English level (LEXTALE, Lemhöfer & Broersma, 2012), and memory span (digit span forward and backward, WAIS-IV, Wechsler, 2008) (see Table 1). Every participant provided written informed consent.

2.2. Stimuli

2.2.1. L1 lexical decision task

2.2.1.1. Targets. Twenty French words (mean number of letters = 6.75, SD = 0.85) were used as targets in the word recognition task. Half of the words were verbs (e.g., cumuler, accumulate), while the others were nouns (e.g., cerisier, cherry tree). All the items were low in frequency (maximum frequency = 13.85 occurrences by million, mean frequency = 3.58 occurrences by million, SD = 3.99). Frequencies were extracted from the book corpus of Lexique 3.80 (New et al., 2001). Target words had either no neighbour or few neighbours (maximum number of neighbours = 3, mean number of neighbours = 1.35, SD = 0.83). Neighbours were always less frequent than the target words (except for the word "brunir", frequency was 0.74, while frequency of its most frequent neighbour "brunie" was 1.01 occurrences by million). The mean frequency of the most frequent neighbours was 2.10 (SD = 2.80). We used low frequency target words with any or few neighbours because given that those words are more difficult to recognize, they are more prone to benefit from both prelexical activation and lexical competition (this choice was also made in other studies as for example: Bowers et al., 2005; Forster & Veres, 1998). The list of stimuli is available in Appendix A.

2.2.1.2. Distractors. Twenty pseudowords (mean number of letters = 6.7, SD = 0.73) were created for the purpose of the lexical decision task by changing one letter from a French word, whose length and frequency were similar to target words. Vowels were always replaced by vowels and consonants by consonants. For seven words the letters were replaced by letters internal to the words, while for the 13 other letters were external. As for the target words previously described, the pseudowords had few neighbours (maximum number of neighbours = 3, mean number of neighbours = 1.25, SD = 0.64, the number of neighbours was not significantly different between words and pseudowords, p value of *t*-test = 0.33) and neighbours were infrequent words (mean frequency of the most frequent neighbours = 3.61, SD = 3.10). The pseudowords created were neither English words, Spanish words nor German words to avoid any confusion. The list of stimuli is available in Appendix A.

Table 1
Mean scores (SD) on background tests for each group and probability associated with Student's paired t-test with two-tailed distribution.

	TE Group	V Group	p
Age	23.64 (3.72)	23.78(3.42)	0.89
Digit span forward	6.75 (1.19)	6.57 (0.79)	0.53
Digit span backward	5.17 (1.09)	5.04 (1.26)	0.72
French LEXTALE percentile	50.68 (33.06)	50.39 (36.41)	0.98
English LEXTALE percentile	71.58 (10.75)	67.90 (8.19)	0.19

2.2.1.3. Primes. Forty new language words (mean number of letters = 6.73, SD = 0.78) were created by changing one letter from the targets previously described. The position of the letter changed was randomly selected (mean position of the letter changed = 4.43, SD = 2.19). Vowels were always replaced by vowels and consonants by consonants. For 18 words the letters were replaced by letters internal to the words, while for the 22 other letters were external. The mean number of French orthographic neighbours was 0.95 (SD = 1.2). The new language words were neither French words, English words, Spanish words nor German words. The French graphotactic constraints were violated in several items in order to make the stimuli like words from a foreign language (for example: cruuger, see the list of stimuli in Appendix B).

The 40 new language words used as primes in the word recognition task were the words to be learned during the learning phase. Half of these new language words were presented to the participants as nouns of a newly created language, the 20 others as verbs. As previously mentioned, these new language words were made to be as close as possible to foreign language words (the French graphotactic constraints were violated in several items, see the list of stimuli in Appendix B).

2.2.2. L2 Word recognition task

In the word recognition task, stimuli included the 40 new language words which had been learned and a set of pseudowords. These pseudowords were created by changing two letters from the new language words to be learned. The decision to change two letters rather than one was taken to avoid having abnormally long response time due to participants who would have analysed each letter of the word to be sure that there is no difference between the learned word and the stimuli presented. The mean number of neighbours was 0.08 (SD = 0.27). The list of distractors is available in Appendix C.

2.3. Procedure

The experiment lasted three days. The first day began with the L1 lexical decision task (session 1) followed by learning phases. The second day concerned only learning phases. The last day started with the L1 lexical decision task (session 2) and ended with a word recognition task in L2. The organisation of tasks is available in Fig. 1. The different phases of the experiment were run under the control of MATLAB software on a Windows laptop. All the tasks were carried out individually in a quiet room.

2.3.1. Learning phases

2.3.1.1. Part 1: presentation of words. The first part of the learning phases was composed of 6 presentations of each item. Presentations were organised by block. Each block was composed of one presentation of each of the 40 items (40 trials by block). Once a block finished, a new block started and so on and so forth until the sixth and last block. Every trial started with the presentation of a visual stimulus for 5000 ms and finished by the visual presentation of the new words referring to the concept depicted by the stimulus (1000 ms, font: Times, size: 50). The intertrial interval was 1000 ms. The type of stimuli presented depended on the learning method. In the video group, stimuli were videos. When the word of the new language was a noun, the video depicted a camera's rotation around a representation of the concept which the word refers to modelled in three dimensions. When the word was a verb, the video

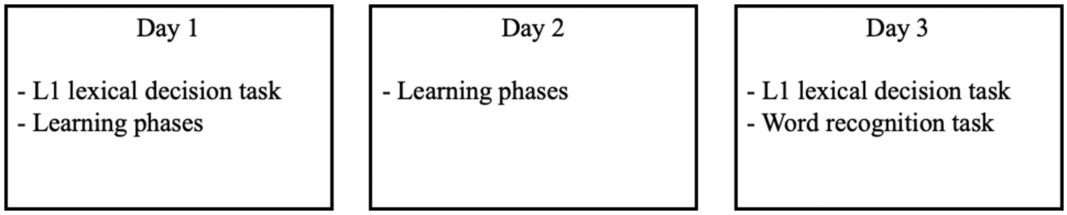

Fig. 1. Organisation of tasks.

represented a man performing the action depicted by the verb. The videos clips were created for this experiment. A screenshot of a video used in the experiment is available in Appendix D. In the translation equivalent (TE) group, stimuli were the French translation equivalents of the L2 words. A handwriting font called "freestyle script" (font size: 50) was used. Participants of both groups received the instruction to learn the associations between the pairs of stimuli.

2.3.1.2. Part 2: semantic retrieval task. In the semantic retrieval task, an L2 word was presented visually (1000 ms) and was followed by three visual stimuli (French words for the translation equivalent group and videos for the video group). The stimuli were arranged vertically one below the other. Among these items, one was associated with the L2 word in the first part of the learning phase (target), while the two others were associated with other words (distractors). The distractors were randomly selected among the learned items. The position of the items was also randomly selected. Participants were instructed to click on the target item as accurately as possible without time constraint (stimuli were presented until participants answered). After every answer, a feedback indicated whether the answer was correct ("correct answer" appearing on the screen on a green background) or not ("incorrect answer" appearing on the screen on a red background). When the answer was incorrect, the failed stimulus was presented again later in the block. Once each of the 40 L2 words was associated with the correct answer, the second and last block started. Thus, each of the 40 L2 words was presented at least once in each block. The number of errors and the number of items failed were recorded by the program in order to make sure that any effect observed in the test phase was not attributable to any difference in the exposure to the L2. The number of errors committed corresponded to the total number of errors committed, e.g. 5 failures on an item equals 5 errors, while, the number of items failed was only the number of items failed, e.g. 5 failures on an item equals 1 item failed. Since neither the number of errors nor the number of items failed were significantly different between the two groups on the first day or the second (see Table 2), the exposure to the L2 was not significantly different.

2.3.2. Test phase

2.3.2.1. L1 lexical decision task. The lexical decision task was composed of 40 trials. Each trial was organised as follows: fixation cross (500 ms), forward mask (500 ms), prime (50 ms) and target word (presented until participants answered). For each trial, participants had to indicate, as rapidly and accurately as possible, whether the stimulus presented was a French word or a pseudoword. Participants answered through the CTRL keys situated on the bottom of the keyboard. The CTRL key situated on the side of the dominant hand was used to indicate that the item presented was a French word, the one situated on the side of the non-dominant hand was used to indicate that the item was a pseudoword. Twenty French words and 20 pseudowords were used in this task. Both were presented either in a related condition or in an unrelated one. In the related condition, the prime and the target differed by one letter, while in the unrelated one, prime and target did not share any letter. To avoid multiple judgments for the same target, the stimuli were counterbalanced across two lists. Stimuli appeared in only one condition in each list (e.g., if an L1 word appeared in List 1 in the related condition, in List 2 it appeared in the unrelated condition). The lists were randomly created. The lexical decision task used on the third day was identical to the one used on the first day.

Table 2
Mean number of errors and mean number of items failed for each group and probability associated with Student's paired t-test with two-tailed distribution.

		TE group	V group	t	p
First day	Errors	18.96 (26.88)	7.78 (13.21)	1.85	0.07
	Items failed	8.60 (9.06)	5.13 (7.87)	1.42	0.16
Second day	Errors	1.84 (2.58)	1.78 (3.32)	0.07	0.95
	Items failed	1.60 (2.14)	1.48 (2.17)	0.2	0.85

2.3.2.2. Word recognition task - L2 words. The word recognition task was composed of 80 trials. Each trial was composed of a fixation cross (500 ms) and a learned new language word or a pseudoword (presented until participants answered). Half of the items presented were learned new language words, the others were pseudowords. Participants had to indicate as accurately and rapidly as possible whether the stimulus presented was a learned item or a pseudoword. Participants answered through the CTRL keys situated on the bottom of the keyboard. The CTRL key situated on the side of the dominant hand was used to indicate that the item presented was a learned word, the one situated on the side of the non-dominant hand was used to indicate that the item was a pseudoword.

3. Results

Data were analysed in the software R (R Core Team, 2017) using a mixed model approach (Baayen et al., 2008) with the lme4 package (Bates, Maechler, et al., 2015). For the random structure we selected a compromise between the maximal approach (e.g. Barr et al., 2013) and the parsimonious approach (e.g. Bates, Kliegl, et al., 2015) by using at a minimum the random effects subject and target (random intercepts) and adding supplementary random effects when the latter improved the model fit. To improve our confidence in the selected models, we used Bayesian statistics. The function brm of the brms package (Bürkner, 2017, 2018; Carpenter et al., 2017) was used to fit the same models and to obtain 95% credible intervals (CrI) as well as posterior distributions for each estimate. The bayes_R2 function of the same package was also used to calculate a Bayesian version of the R^2 (Gelman et al., 2019).

3.1. L1 lexical decision task

Three L1 words were removed from the analysis because they were poorly known, for example, the percentage of correct responses of brunir (69%, SD: 47; in the first session: 71%, SD: 46) was lower than that of the other items (mean percentage of correct responses to the 20 items: 90%, SD: 29; in the first session: 90%, standard deviation: 30). As L1 words serve as probes for lexical competition, one should make sure that they are actually in the lexicon. These words being unknown by several participants (not integrated in their lexicon), it was not possible to use them to test further lexical competition.

3.1.1. Accuracy

Given that accuracy is a dichotomous variable, errors were analysed with binomial mixed models. To select the best model, a backward elimination procedure was used. The model selection procedure started with a complete model including the 2 following random effects (participants and targets) and these two fixed effect factors and their interaction (Session: before learning phase, i.e. session 1, after learning phase, i.e. session 2; Learning method: VE, TE). According to the Akaike Information Criterion (AIC), the best model included no fixed effect factor. For completeness, the results of the parameters in the complete model and the output from the Bayesian analysis are reported in Table 3. For information, the mean percentage of correct responses in the video group was 95% (standard deviation: 22; 93% for detection of words, standard deviation: 25) and was 93% in the translation equivalent group (standard deviation: 25, 91% for detection of words, standard deviation: 28).

3.1.2. Response times

Incorrect responses and pseudowords were not included in the analyses. A graphical inspection based on boxplot representation indicated that RT lower than 100 ms and greater than 1300 ms should be considered as outliers.

To select the best model, an automatic backward elimination procedure using the function step from the lmerTest package (Kuznetsova et al., 2017) was used. The model selection procedure started with a complete model including the 3 following random effects (participant, target and list) and three fixed effect factors and their interactions (Session: before learning phase, i.e. session 1, after learning phase, i.e. session 2; Learning method: VE, TE; Relatedness of the prime: related, unrelated). The best model was the one with the following formula: RT ~ Relatedness + Learning Method + Session + Group:Session + (1 | Participant) + (1 | Target) + (1 | List). The results of the parameters in this model and the output from the Bayesian analysis are reported in Table 4. The Bayesian model seemed to correctly fit our data (Bayes $R^2 = 0.422$ (SE = 0.016, CrI = [0.388; 0.453]).

The fact that the best model included an interaction effect between learning method and session indicated that the patterns were significantly different according to the session and the learning method. Therefore, we analysed separately the data from the two sessions. In the first session, the model with the best fit was the one with only the effect of relatedness of the prime ($\chi2 = 6.72$, $p < .01$) see Table 5. While, in the second session, it was the model with the interaction between learning method and relatedness between prime and target ($\chi2 = 5.15$, $p = .023$), see Table 6.

The best model for the second session showed that the patterns of response times were significantly different between the two groups in this session. Indeed, in the second session, the model with relatedness between prime and target significantly improved the model fit for the translation equivalent group (see Table 7) but not for the video group (see Table 8). In other words, there was a significant priming effect in the translation equivalent group but not in the video group. Mean response times are reported in Fig. 2 (there are also available in Appendix E).

3.2. Word recognition task - L2 words

3.2.1. Accuracy

Given that accuracy is a dichotomous variable, errors were analysed with binomial mixed models. We compared a model including the effect of learning method (VE, TE) to a base model (without fixed effect factors). The models included a random effect factor (subject, adding the factor "target" leads to convergence problem).

As expected, the model including learning method did not improve the fit of the base model (for word recognition: $\chi2 = 0.34$, $p = .56$; for word recognition and rejection of pseudowords: $\chi2 = 1.41$, $p = .23$). For completeness, the results of the parameters in the complete model and the output from the Bayesian analysis are reported in Table 9 for word recognition and in Table 10 for both word recognition and rejection of pseudowords. The mean percentages of correct responses are shown in Table 11.

3.2.2. Response times

Incorrect responses were not included in the analyses. Reaction times longer than 2000 ms and shorter than 100 ms were also excluded (3% of the data). To select the best model, an automatic backward elimination procedure was used with the function step from the lmerTest package (Kuznetsova et al., 2017). The model selection procedure started with a complete model including the 2 following random effects (participant, target) and the fixed effect factor learning method (VE, TE).

As expected, the model including the effect of learning method did not improve the fit of the base model (for word recognition: $\chi^2 = 0.13$, $p = .71$; for word recognition and rejection of pseudowords: $\chi^2 = 0.01$, $p = .99$). The mean response times are available in Table 11.

4. Discussion

The aim of this experiment was to study the integration of L2 words into the mental lexicon and to investigate the effect of learning method on this process with a paradigm using two learning methods: translation equivalent and video.

To investigate the integration of L2 words, we used the PLE (Forster & Veres, 1998, see also Davis & Lupker, 2006) which is based on form priming (Forster & Davis, 1984). As a reminder, primes and targets in form priming share all but one letter in the related condition but no letter in the unrelated condition. In this paradigm, the priming effect depends of the lexical status of the prime. When the prime is a pseudoword, the form overlap facilitates word recognition. However, when the prime is a word, the effect is reduced, vanishes or becomes inhibitory because words used as primes activate their lexical representations, leading to the inhibition of the representations orthographically close and to the absence of facilitation effect for word detection.

In accordance with previous experiments, results of the first session showed that priming by pseudowords leads to a facilitation for the detection of L1 words in the related condition in comparison with the unrelated condition for the two groups of participants (video and translation equivalent). After learning, the pattern of results depended on the learning method. In the video group, L2 words (pseudowords learned as new language words) no longer facilitated L1 word recognition in the related condition as compared with the unrelated condition. However, in the translation equivalent group, L2 words (pseudowords learned as L2 words) still facilitated L1 word recognition in the related

Table 3
Summary of the complete model for accuracy (and in italics output from the Bayesian analysis).

Predictors	b	SE b	z	p	*b*	*SE b*	*95%CrI*
(Intercept)	2.886	0.239	12.082	<0.001	*2.92*	*0.25*	*[2.45 ; 3.46]*
Session	0.348	0.186	1.874	0.061	*0.35*	*0.19*	*[−0.02 ; 0.73]*
Learning Method	0.378	0.295	1.281	0.200	*0.38*	*0.31*	*[−0.22 ; 1.01]*
Session x Learning Method	−0.171	0.285	−0.601	0.548	*−0.17*	*0.30*	*[−0.76 ; 0.41]*

Table 4
Summary of the first model for response times (and in italics output from the Bayesian analysis).

Predictors	b	SE b	t	p	*b*	*SE b*	*95% CrI*
(Intercept)	769.831	45.165	17.045	<0.001	*−15.51*	*16.21*	*[−47.79; 16.03]*
Relatedness	22.256	8.169	2.724	<0.01	*22.23*	*5.88*	*[10.66; 33.91]*
Learning method	5.035	34.674	0.145	0.885	*5.79*	*24.53*	*[−42.81; 51.73]*
Session	−12.975	11.324	−1.146	0.252	*−13.23*	*7.33*	*[−27.60; 1.27]*
Learning method x Session	36.317	16.167	2.246	<0.05	*36.51*	*10.54*	*[15.64; 57.58]*

Table 5
Summary of the model for response times in the first session (and in italics output from the Bayesian analysis).

Predictors	b	SE b	t	p	*b*	*SE b*	*95%CrI*
(Intercept)	768.917	37.211	20.664	<0.001	*−13.28*	*10.81*	*[−35.05; 7.33]*
Relatedness	27.874	10.714	2.602	<0.01	*27.57*	*7.60*	*[12.72; 42.33]*

Bayes_R^2 = 0.475 (SE = 0.022, IC = [0.429; 0.516].

Table 6
Summary of the model for response times in the second session (and output from the Bayesian analysis).

Predictors	b	SE b	t	p	*b*	*SE b*	*95%CrI*
(Intercept)	752.752	47.372	15.890	<0.001	*−35.46*	*16.73*	*[−68.07; −2.39]*
Relatedness	40.098	15.696	2.555	<0.05	*40.22*	*8.51*	*[23.68; 57.09]*
Learning method	66.778	40.930	1.632	0.109	*50.58*	*26.26*	*[−2.20; 98.04]*
Relatedness x Learning method	−51.973	22.825	−2.277	<0.05	*−37.58*	*12.37*	*[−61.67; −12.77]*

Bayes_R^2 = 0.422 (SE 0.016, IC = [0.390; 0.453].

Table 7
Summary of the model for response times in the second session for the translation equivalent group (and in italics output from the Bayesian analysis).

Predictors	b	SE b	t	p	*b*	*SE b*	*95%CI*
(Intercept)	755.303	39.180	19.278	<0.001	*−7.80*	*10.91*	*[−29.37; 13.60]*
Relatedness	36.500	10.915	3.344	<0.001	*16.23*	*7.32*	*[2.00; 30.44]*

Bayes_R^2 = 0.467 (SE = 0.022, IC = [0.423; 0.507].

Table 8
Summary of the model for response times in the second session for the video group (and in italics output from the Bayesian analysis).

Predictors	b	SE b	t	p	*b*	*SE b*	*95%CI*
(Intercept)	796.285	70.771	11.252	<0.05	*−0.60*	*10.73*	*[−20.93; 20.84]*
Relatedness	5.496	12.285	0.447	0.655	*2.17*	*7.57*	*[−12.55; 17.16]*

Bayes_R^2 = 0.361 (SE = 0.025, IC = [0.309; 0.409].

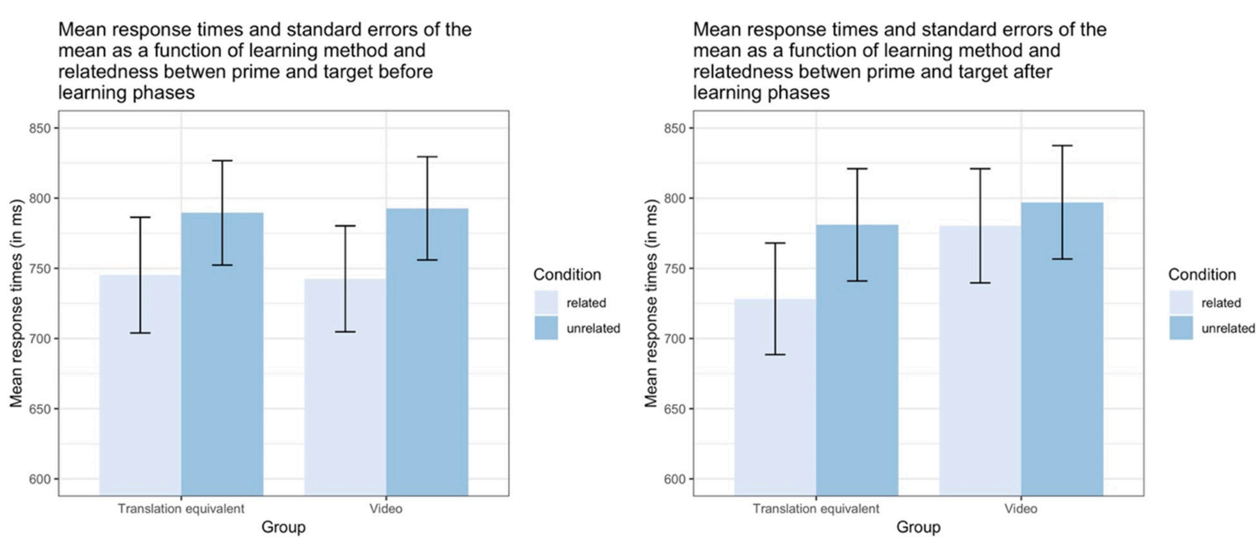

Fig. 2. Reaction times as a function of learning method and relatedness between prime and target before and after learning phases.

Table 9
Summary of the complete model for accuracy for word recognition (and in italics output from the Bayesian analysis).

Predictors	b	SE b	z	p	*b*	*SE b*	*95%CI*
(Intercept)	2.853	0.321	8.878	<0.001	*2.87*	*0.38*	*[2.15; 3.62]*
Learning method	0.272	0.460	0.591	0.555	*0.32*	*0.53*	*[−0.73; 1.38]*

Table 10
Summary of the model for accuracy for both word recognition and rejection of pseudowords and in italics output from the Bayesian analysis).

Predictors	b	SE b	z	p	*b*	*SE b*	*95%CI*
(Intercept)	3.073	0.260	11.807	<0.001	*3.11*	*0.27*	*[2.59; 3.65]*
Learning method	0.455	0.376	1.211	0.226	*0.44*	*0.40*	*[−0.35; 1.20]*

Table 11
Mean reaction times (SD) and percentage of correct responses as a function of learning method in the word recognition task.

	Translation equivalent group	Video group
Mean percentage of correct responses (SD)	92 (28)	96 (20)
Mean percentage of correct word recognition (SD)	90 (30)	93 (25)
Mean reaction times for correct responses (SD)	763 (260)	767 (250)
Mean reaction times for correct word recognition (SD)	751 (260)	770 (257)

condition as compared with the unrelated one. As we will discuss in the next paragraphs, there were two main results. Firstly, there was a cross language PLE since the effect of primes on L1 words in the video group was modulated by the lexical status of the primes. Secondly, there was an effect of learning method on the integration process into the mental lexicon since the absence of a priming effect was observed only with the video group.

Previous research has already shown that a PLE may be observed with newly learned words both in L1 (Qiao & Forster, 2013) and in L2 (Elgort, 2011; Elgort & Piasecki, 2014). Nevertheless, to our knowledge, no previous study has reported any cross-language PLE. Our results clearly indicate for the first-time a cross-language PLE when L2 words serve as primes and L1 words as targets. The implications of these results are in accordance with the findings of Elgort (2011) and Elgort and Piasecki (2014), pleading in favour of a lexicon shared by the two languages. On the contrary, the cross-language PLE is in opposition with the results and interpretations of Qiao and Forster (2017). They posited that different memory systems are involved for L1 and L2 words when the latter are learned after a critical period (see also: Jiang & Forster, 2001). As previously mentioned, the results of Qiao and Forster (2017) might be attributable to participant language' characteristics. As a reminder, the two languages of the participants did not have the same writing system: the L1 of the participants had a logographic writing system (Chinese), while the second one had an alphabetic one (English). However, the heterogeneity in the results is not attributable to differences in presentation duration. Although the differences in presentation duration in the experiments of Elgort (522 ms in Elgort, 2011; 490 ms in Elgort & Piasecki, 2014) and Qiao & Forster (2017, 50ms) might have led to differences in the results, this cannot be the case for our experiment since durations were identical to those used in Qiao and Forster (2017), i.e. 50 ms. Using an L2 (pseudowords presented as foreign language words, this point will be discussed later, in the part devoted to the limitations of the experiment) which shared the same writing system as the L1 and a short presentation duration (50 ms), our study suggests that new L2 lexical representations interact with L1 representations.

As previously indicated, our second result concerns the effect of teaching method. A PLE was observed only with the video method. It is worth noting that the L2 word recognition task did not reveal any differences between the two groups, so the differences in L2 words integration cannot be explained by differences in L2 word knowledge. As a reminder, L2 word knowledge was reinforced through a semantic retrieval task in which a 100% criterion was used to promote learning without excluding any participants. This criterion could have led to the differences observed between the two groups, since one group may have been more exposed to the L2 words than the other one. Nevertheless, the analysis showed that it was not the case since neither the number of errors, nor the number of items failed were significantly different between the two groups. Therefore, results on the L1 lexical decision task showed a) that L2 words were swiftly integrated with already existing representations when they were learned with a video method allowing a soft immersion in an L2 environment and/or b) that a learning method based on translation equivalent delays this integration.

Results also suggest that learning vocabulary through interactions with concepts led to the integration of L2 words into a lexicon shared by the two languages. Nevertheless, it is not possible to conclude whether learning new language words with translation equivalent leads to store L2 words in a different lexicon of L1 words or if this learning method did not allow L2 words to participate in the lexical competition process in the first steps of learning. A parallel can be drawn with the assumption of BIA-d (Grainger et al., 2010) namely that L2 words are stored in different lexicons in novice learners while there is a single lexicon for the two languages in proficient learners. Indeed, our results suggest that learning new words through interaction with concepts led to integrate L1 words and L2 words into the same lexicon, while learning through translation equivalent did not in the first steps of learning. Besides the impact of proficiency on word integration that is assumed by BIA-d, our study shows that the learning method also impacts L2 word integration.

The present study has some limitations. First, the decision to use pseudowords rather than new language words to control some factors including L1 neighbourhood can arouse questions, especially about the strategy used by the participant to learn these L2 words. Are these pseudowords learned through a strategy normally use to learn new L1 words? If so, are the conclusions drawn for L1 learning rather than for L2 learning? In light of the choice made in this experiment, this interpretation is unlikely for several reasons. First, an important point is that we told to our participants that they were learning words from a new language. Indeed, as pointed out by Meade et al. (2018) this instruction allows to "set a specific learning context, which has been shown to be an important factor in numerous previous studies of L2 acquisition (see, e.g., Collentine & Freed, 2004, for a review)". Secondly, the type of explicit learning used, i.e. learning new language words for familiar concepts that already had clear L1 labels, is a common way to learn new L2 words for adults, while L1 words are generally learned in a more implicit way as for example during reading (for a discussion see Meade et al., 2018). Furthermore, this was reinforced by the fact that our new language words violated the French graphotactic constraints in several items. It should also be note that pseudoword learning is usual in L2 studies using a learning paradigm and the PLE (see for example: Elgort, 2011; Elgort & Piasecki, 2014; Qiao et al., 2009; Qiao & Forster, 2013; Qiao & Forster, 2017).

Another limitation is the short time period of the experiment. Further research is needed to investigate the integration of L2 words in the translation equivalent group. Nevertheless, the present findings already demonstrate that learning methods modulate L2 word integration.

Finally, and related to the latter point, our participants only learned 20 new words, and among them only 17 could be analysed from the PLE effects due to lack of knowledge of the L1 word which should act as the word to be competed. While Bayesian tests could display the strength of the results, further studies should involve more words to be learned, and therefore more learning sessions.

In summary, this study shows a cross-language PLE with newly learned L2 words used as primes and L1 words as targets. It also demonstrates the effect of learning method in this process since method allowing a soft immersion in an L2 environment promoted the integration of L2 words into the mental lexicon.

CRediT authorship contribution statement

Gary Boddaert: Conceptualisation, Methodology, Software, Validation, Formal Analysis, Investigation, Writing – Original Draft, Writing – Review & Editing. **Camille Cornut:** Validation, Formal Analysis, Writing – Review & Editing. **Séverine Casalis:** Conceptualisation, Methodology, Validation, Formal Analysis, Writing – Original Draft, Writing – Review & Editing, Supervision.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Acknowledgments

We thank Perle Aissati for help in collecting data and Laurent Ott for help in preparing the experiment.

Appendix A. Stimuli used in lexical decision task and their characteristics (L1 words)

Item	Lexical status	Number of letters	Word frequency in French	Number of French orthographic neighbour
Anonymat	Word	8	3.04	0
Assagir	Word	7	0.41	3
Assainir	Word	8	0.68	2
Bambin	Word	6	2.16	2
Brunir	Word	6	0.74	3
Cerisier	Word	8	1.49	1
Cordon	Word	6	9.19	2
Cumuler	Word	7	0.2	1
Dicter	Word	6	1.89	2
Donation	Word	8	0.54	1
Expulser	Word	8	2.3	1
Flacon	Word	6	11.82	1
Grincer	Word	7	4.46	2
Jasmin	Word	6	4.19	1
Jungle	Word	6	7.43	2
Lavabo	Word	6	13.85	0
Mincir	Word	6	0	1
Muguet	Word	6	3.85	0
Polluer	Word	7	0.34	2
Stopper	Word	7	3.11	3
Armule	Pseudoword	6	Pseudoword	1
Atrosoir	Pseudoword	8	Pseudoword	1
Bascotte	Pseudoword	8	Pseudoword	3
Brebas	Pseudoword	6	Pseudoword	1
Chutur	Pseudoword	6	Pseudoword	1
Clignir	Pseudoword	7	Pseudoword	1
Confoi	Pseudoword	6	Pseudoword	1
Crouger	Pseudoword	7	Pseudoword	1
Divirger	Pseudoword	8	Pseudoword	1
Empacer	Pseudoword	7	Pseudoword	3
Exclute	Pseudoword	7	Pseudoword	1
Flirmer	Pseudoword	7	Pseudoword	1
Jaguir	Pseudoword	6	Pseudoword	1
Matalot	Pseudoword	7	Pseudoword	1
Sobbet	Pseudoword	6	Pseudoword	1
Soumon	Pseudoword	6	Pseudoword	2
Spogan	Pseudoword	6	Pseudoword	1
Tutoyar	Pseudoword	7	Pseudoword	1
Unduler	Pseudoword	7	Pseudoword	1
Urgile	Pseudoword	6	Pseudoword	1

Appendix B. Learned words (i.e. primes) and L1 translation equivalents

Learned word (i.e. prime)	Translation equivalent in L1	Number of letters of the learned word	Word frequency of the L1 translation equivalent	Number of French orthographic neighbours in L1
Assagif	Sculpter	7	2.16	4
Assainin	Scier	8	2.43	3
Atrisoir	Sous-marin	8	3.98	1
Bamban	Stéthoscope	6	0.54	1
Bascotto	Montagne	8	49.8	0
Brabas	Cactus	6	2.3	0
Brenir	Écrire	6	116.15	1
Cerusier	Bague	8	16.15	1

(continued on next page)

(*continued*)

Learned word (i.e. prime)	Translation equivalent in L1	Number of letters of the learned word	Word frequency of the L1 translation equivalent	Number of French orthographic neighbours in L1
Chutut	Tricoter	6	3.11	0
Clignin	Arroser	7	4.46	0
Cruuger	Plier	7	10.68	0
Cumulet	Percer	7	11.22	2
Devirger	Coller	8	12.36	0
Docter	Lacer	6	0.68	4
Donition	Hérisson	8	1.76	2
Empacem	Découper	7	4.93	0
Ermule	Champignon	6	3.99	0
Exclote	Visser	7	0.88	0
Expulter	Repasser	8	9.39	1
Fladon	Trèfle	6	4.19	1
Flirmir	Signer	7	13.51	0
Gonfoi	Lit	6	340.6	0
Griscer	Manger	7	138.31	1
Jammin	Aigle	6	7.91	1
Jeguir	Coudre	6	8.65	0
Jengle	Raisin	6	4.86	3
Jordon	Cerf-volant	6	1.22	1
Lavabi	Étoile	6	31.02	1
Matalat	Enveloppe	7	33.11	0
Mincin	Boire	6	102.3	2
Polluet	Peindre	7	22.64	3
Puguet	Fraise	6	3.99	1
Sobbit	Avion	6	46.82	0
Soubon	Nuage	6	26.49	0
Spotan	Aquarium	6	5.2	0
Stoppem	Pêcher	7	6.35	3
Tutoyad	Tondre	7	1.01	0
Uggile	Hélicoptère	6	2.43	0
Undulir	Dessiner	7	12.97	0
Unonymat	Hamburger	8	0.41	1

Appendix C. Distractors used in the L2 word recognition task

Distractor	Number of letters	Number of French orthographic neighbours
Absailin	8	0
Aloppem	7	0
Arpacem	7	0
Astigif	7	0
Atrisuin	8	0
Basconti	8	0
Bengre	6	0
Bracir	6	0
Butulet	7	0
Carusiet	8	0
Clatut	6	0
Clignum	7	0
Demirgem	8	0
Donitius	8	0
Donper	6	1
Doubor	6	0
Druugor	7	0
Elanon	6	0
Erdale	6	0
Exchota	7	0
Expirter	8	1
Gerfoi	6	0
Ghirmir	7	0
Glabas	6	0
Grascir	7	0
Jegoar	6	0
Jordut	6	0
Lavamu	6	0
Matamar	7	0
Mincat	6	0
Pobluem	7	0
Puglit	6	0
Romban	6	0
Spitat	6	0
Subrit	6	0
Tammon	6	1
Totoyac	7	0

(*continued on next page*)

(*continued*)

Distractor	Number of letters	Number of French orthographic neighbours
Uggame	6	0
Undutis	7	0
Unonaman	8	0

Appendix D. Screenshot of video used in learning phases

Appendix E. Mean reaction times (SD) in lexical decision task as a function of session, learning method and relatedness between prime and target

First session				Second session			
Video method		Translation equivalent method		Video method		Translation equivalent method	
Related	Unrelated	Related	Unrelated	Related	Unrelated	Related	Unrelated
743 (185)	793 (180)	745 (202)	790 (182)	780 (199)	797 (198)	728 (195)	781 (196)

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. Retrieved from http://arxiv.org/abs/1506.04967.

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

Bijeljac-babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory & Cognition, 25*, 447–457. https://doi.org/10.3758/BF03201121

Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Interfering neighbours: The impact of novel word learning on the identification of visually similar words. *Cognition, 97*(3), B45–B54. https://doi.org/10.1016/j.cognition.2005.02.002

Brysbaert, M. (2013). LEXTALE_FR: A fast, free, and efficient test to measure language proficiency in French. *Psychologica Belgica, 53*, 23–37.

Bürkner, P. C. (2017). An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software, 80*(1). https://doi.org/10.18637/jss.v080.i01

Bürkner, P. C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal, 10*(1), 395–411. https://doi.org/10.32614/RJ-2018-017

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., … Ridell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*. https://doi.org/10.18637/jss.v076.i01

Collentine, J., & Freed, B. F. (2004). Learning context and its effects on second language acquisition. *Studies in Second Language Acquisition, 26*, 153–171.

Comesaña, M., Perea, M., Piñeiro, A., & Fraga, I. (2009). Vocabulary teaching strategies and conceptual representations of words in L2 in children: Evidence with novice learners. *Journal of Experimental Child Psychology, 104*, 22–33. https://doi.org/10.1016/j.jecp.2008.10.004

Comesaña, M., Soares, A. P., & Lima, C. (2010). Semantic representations of new cognate vs. noncognate words: Evidence from two second language learning methods. *Procedia - Social and Behavioral Sciences, 5*, 199–203. https://doi.org/10.1016/j.sbspro.2010.07.072

Comesaña, M., Soares, A. P., Sánchez-Casas, R., & Lima, C. (2012). Lexical and semantic representations in the acquisition of L2 cognate and non-cognate words: Evidence from two learning methods in children. *British Journal of Psychology, 103*, 378–392. https://doi.org/10.1111/j.2044-8295.2011.02080.x

Davis, C. J., & Lupker, S. J. (2006). Masked inhibitory priming in English: Evidence for lexical inhibition. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 668–687. https://doi.org/10.1037/0096-1523.32.3.668

Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: Neural and behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences, 364*, 3773–3800. https://doi.org/10.1098/rstb.2009.0111

Dijkstra, T., & van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5*, 175–197. https://doi.org/10.1017/S1366728902003012

Dijkstra, T., Wahl, A., Buytenhuijs, F., Van Halem, N., Al-Jibouri, Z., De Korte, M., & Rekké, S. (2018). *Multilink: A computational model for bilingual word recognition and word translation* (pp. 1–23). Language and Cognition: Bilingualism. https://doi.org/10.1017/S1366728918000287

Dong, Y., Gui, S., & Macwhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition, 8*(3), 221–238. https://doi.org/10.1017/S1366728905002270

Dumay, N., & Gaskell, M. G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18*(1), 35–39. https://doi.org/10.1111/j.1467- 9280.2007.01845.x

Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning, 61*, 367–413. https://doi.org/10.1111/j.1467-9922.2010.00613.x

Elgort, I., & Piasecki, A. E. (2014). The effect of a bilingual learning mode on the establishment of lexical semantic representations in the L2. *Bilingualism: Language and Cognition, 17*(3), 572–588. https://doi.org/10.1017/S1366728913000588

Forster, K. I., & Davis, C. (1984). Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 10*, 680–698. https://doi.org/10.1037/0278-7393.10.4.680

Forster, K. I., & Veres, C. (1998). The prime lexicality effect: Form-priming as a function of prime awareness, lexical status, and discrimination difficulty. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 498–514. https://doi.org/10.1037//0278-7393.24.2.498

Gaskell, M., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition, 89*, 105–132. https://doi.org/10.1016/S0010-0277(03)00070-2

Gelman, A., Goodrich, B., Gabry, J., & Vehtari, A. (2019). R-squared for Bayesian regression models. *The American Statistician, 73*(3), 307–309. https://doi.org/10.1080/00031305.2018.1549100

Grainger, J., Midgley, K., & Holcomb, P. J. (2010). Chapter 14. Re-thinking the bilingual interactive-activation model from a developmental perspective (BIA-d). In M. Kail, & M. Hickmann (Eds.), *Language acquisition and language disorders* (pp. 267–283). Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/lald.52.18gra.

Jiang, N., & Forster, K. I. (2001). Cross-language priming asymmetries in lexical decision and episodic recognition. *Journal of Memory and Language, 44*(1), 32–51. https://doi.org/10.1006/jmla.2000.2737

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149–174. https://doi.org/10.1006/jmla.1994.1008

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*, 1–26. https://doi.org/10.18637/jss.v082.i13

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods, 44*, 325–343. https://doi.org/10.3758/s13428-011-0146-0

Lindsay, S., & Gaskell, M. G. (2013). Lexical integration of novel words without sleep. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*(2), 608–622. https://doi.org/10.1037/a0029243

Meade, G., & Dijkstra, T. (2017). Mechanisms underlying word learning in second language acquisition. In M. Libben, M. Goral, & G. Libben (Eds.), *Bilingual Processing and Acquisition (Vol. 6, p. 49–72)*. John Benjamins Publishing Company. https://doi.org/10.1075/bpa.6.03mea.

Meade, G., Midgley, K. J., Dijkstra, T., & Holcomb, P. J. (2018). Cross-language neighborhood effects in learners indicative of an integrated lexicon. *Journal of Cognitive Neuroscience, 30*(1), 70–85. https://doi.org/10.1162/jocn_a_01184

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet: LEXIQUE. *L'Année Psychologique, 101*, 447–462. http://www.lexique.org.

Pavlenko, A. (2009). The bilingual mental lexicon: Interdisciplinary approaches. In *Bristol, UK*. Buffalo, NY: Multilingual Matters.

Qiao, X., & Forster, K. I. (2013). Novel word lexicalization and the prime lexicality effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 1064–1074. https://doi.org/10.1037/a0030528

Qiao, X., & Forster, K. I. (2017). Is the L2 lexicon different from the L1 lexicon? Evidence from novel word lexicalization. *Cognition, 158*, 147–152. https://doi.org/10.1016/j.cognition.2016.10.026

Qiao, X., Forster, K. I., & Witzel, N. (2009). Is banara really a word? *Cognition, 113*, 254–257. https://doi.org/10.1016/j.cognition.2009.08.006

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.

Wang, H.-C., Savage, G., Gaskell, M. G., Paulin, T., Robidoux, S., & Castles, A. (2017). Bedding down new words: Sleep promotes the emergence of lexical competition in visual word recognition. *Psychonomic Bulletin & Review, 24*(4), 1186–1193. https://doi.org/10.3758/s13423-016-1182-7

Wechsler, D. (2008). *Wechsler Adult Intelligence Scale-Fourth Edition: Technical and interpretative manual*. San Antonio, TX: Pearson Assessment.