

CORNUT, Camille, MAHÉ, Gwendoline, et CASALIS, Séverine. L2 word recognition in French–English late bilinguals: does modality matter?. *Bilingualism: Language and Cognition*, 2021, p. 1-16.

Short title: Modality effect in L2 word recognition.

Complete title: L2 word recognition in French-English late bilinguals: does modality matter?

Authors:

Camille CORNUT<sup>1</sup>, Gwendoline MAHE<sup>1</sup>, Séverine CASALIS<sup>1</sup>

<sup>1</sup> Univ. Lille, CNRS, UMR 9193 - SCALab - Sciences Cognitives et Sciences Affectives, F-59000 Lille, France

\* Acknowledgements:

The authors are grateful for the region Hauts-de-France and the University of Lille for the grant that made this study possible. The study benefits also for the project APPREL2, grant ANR (National Research Agency)-16-CE28-0009-01 awarded to Séverine Casalis.

We thank also Marc Brysbaert and the two anonymous reviewers for their helpful comments to improve the paper.

Address for correspondence:

Pr. Séverine CASALIS

severine.casalis@univ-lille.fr

+33 3 20 41 63 69

Université de Lille

Laboratoire SCALab

Domaine Universitaire Pont de Bois

Rue du Barreau

BP60149

59653 Villeneuve d'Ascq

France

**Abstract:**

Research in second language (L2) learning often considers one modality only during task completion. It is unclear if L2 performance is as accurate whatever the modality. L2 learning at school is characterized by a predominance of written materials. One might expect written L2 word recognition to be more accurate than spoken one. This modality effect could also depend on L2 proficiency and the presence of cognate items, closer orthographically than phonologically for most language pairs. Two experiments were conducted with 50 intermediate proficiency French-English bilinguals. Experiment 1 highlighted this modality effect on accuracy and a session effect reflecting a benefit from oral to written modality on latency. In Experiment 2, which included both cognate and non-cognate words, modality effect was even stronger for cognate words and cognate effect depended on modality. In both experiments, these effects depend on L2 proficiency. These findings are discussed according to bilingual word recognition models.

**Keywords:** bilingualism, word recognition, modality effect, lexical representations

## **Introduction.**

In today's globalized world, being proficient in a foreign language has become crucial for most people. For this reason, second language (L2) teaching is now a major challenge for public policies. Nevertheless, even if language fluency is based on both oral and written communication skills, pupils are mainly exposed to written materials due to the practicalities of L2 learning at school. From a theoretical point of view, most L2 word recognition models consider word recognition processes in one modality only, either oral or written. The question arises, however, about the issue of the impact of modality on L2 word recognition, especially among individuals who have not yet reached a high level of L2 proficiency. In the present study, we therefore examined the impact of modality on L2 word recognition among intermediate proficiency bilinguals. This issue was investigated both for translation equivalents that do not share any formal overlap between languages (non-cognate words) and for translation equivalents that share orthographic overlap in mother tongue (L1) and L2 (cognate words).

Given the predominance of written materials in learning at school, we first considered written word recognition. The dual-route model hypothesizes two possible pathways for this process in L1: phonological and lexical pathways (Coltheart et al., 2001). Grapheme-phoneme conversion (GPC) rules are the basis of the phonological pathway in which readers use these rules to decode pseudowords and words they have never encountered before. The efficacy of decoding depends on the consistency of the orthography of the language considered: the less the orthography is consistent, the more difficult it is for readers to correctly decode a new word and/or a pseudoword, GPC rules being less systematic. On the contrary, the lexical pathway is based on a direct access to the orthographic form of familiar words already integrated in the mental lexicon of the reader, secondarily activating the associated phonological form.

Considering L2 word learning, questions arise in L2 learners whether and when word recognition mechanisms and the lexicon are shared or not between L1 and L2 during L2 acquisition. Interestingly, L2 models (e.g., BIA - Dijkstra, Van Heuven, & Grainger, 1998; BIA+ - Dijkstra & van Heuven, 2002) consider mostly high proficiency bilinguals, even if the recent model Multilink includes a component related to L2 proficiency (Dijkstra et al., 2019) and if BIA (Dijkstra et al., 1998) simulated L2 proficiency variations. Note that the bilingual interactive-activation model – developmental (BIA-d: Grainger, Midgley, & Holcomb, 2010) was designed to describe the development of lexical knowledge among late learners of L2 in written form. In this model, the organization of relations between the lexical forms of the two languages and the semantic level depends on L2 mastery. Among less proficiency bilinguals, the BIA-d assumes an indirect access to meaning via translation equivalents. Importantly, this access is especially facilitated by the existence of particular words that share both semantic and formal aspects. These words, called cognate words (e.g., in French and English: *guide*), are particularly relevant in indicating how L1 and L2 interact, or in other words how L1 is involved in L2 processing. The Multilink model includes both lexico-semantic processing and word comprehension and production (Dijkstra et al., 2019). To do so, it simulates word recognition and production during the most common psycholinguistic tasks, considering both cognate and non-cognate words and taking L2 proficiency into account. It assumes a shared network between languages, allowing the cognate facilitation effect in recognition by co-activation of the shared meaning of words in both languages.

Many studies with different paradigms have highlighted this cognate facilitation effect among bilingual expert readers, cognate words in participants' L1 eliciting shorter response latencies than non-cognate words (see for a meta-analytic review Lauro & Schwartz, 2017; see for evidence in ERP study: Yudes, Macizo, & Bajo, 2010). Moreover, low proficiency bilinguals showed a greater cognate effect in their non-dominant language, whereas high proficiency bilinguals presented a similar cognate effect in both languages (see for a meta-analytic review

Lauro & Schwartz, 2017; see for evidence in eye-tracking study: Pivneva, Mercier, & Titone, 2014). Yet less is known about cognate effect among the particular population of intermediate proficiency late learners of L2 in written form. Additionally, this cognate effect has been studied mostly in written modality. Indeed, the overlap between two languages sharing the same alphabetic system is more often based on spelling than on phonology, as is the case between French and English (see for example *guide* which is spelled in the same way but pronounced differently). Thus, one may wonder how this cognate effect manifests itself in oral modality among intermediate proficiency bilinguals, as L2 words are learned mostly in written form at school (see for cognate effect considering L2 words learned in oral modality Valente et al., 2018, see also for cognate effect in oral modality among high-proficiency bilinguals Wu et al., 2019).

Despite the large amount of work devoted to high-proficiency bilingual word processing, less is known about intermediate proficiency processing in L2 learners who have been studying the L2 in a classroom setting. For this reason, in the current study, we investigated word recognition in intermediate proficiency bilinguals characterized by a limited number of words in their L2 orthographic lexicon. Moreover, L2 learning in a school context in France, being characterized by a low exposure to the L2 and by a predominance of written materials, those learners have less auditory than visual exposure in L2. Consequently, they have less occasion to create phonological representations of L2 words. Thus, their limited access to the phonological form of written words in L2 increases the need to use orthography-to-phonology mappings to decode the words they encounter. On the other hand, their level of proficiency in L2 does not allow them to be certain of the correct mappings to use (see Figure 1 for an example of this difficulty). This is particularly true when the L2 is English, a language with an inconsistent orthography. This inconsistency may cause L2 learners to be uncertain of the correct mappings between orthography and phonology. Here, we focus on learners who have a less inconsistent orthography in their L1 - French, than in their L2 - English (for the difficulty to learn orthography-to-phonology mappings of an L2 with a more

inconsistent orthography than that of L1, see Ziegler & Goswami, 2005). Moreover, the existence of an incongruence between French and English (*i.e.*, the existence of both phonemes and graphemes specific of each language, and the existence of common graphemes corresponding to different phonemes in both languages) makes it difficult for French learners to learn the specific orthography-to-phonology mappings of English, these mappings being different in both languages. To determine whether L1 and/or L2 rules are activated during L2 written word recognition, Commissaire, Duncan and Casalis (2019) proposed English lexical decision tasks to French adolescents during their first or third year of formal English learning. They included pseudo-homophones in their lists of stimuli that sounded like L2 real words both when using L1 and L2 orthography-to-phonology mappings on the one hand, and pseudo-homophones that sounded like L1 real words when using L2 orthography-to-phonology mappings on the other. They found pseudo-homophone interferences in all conditions, suggesting an automatic activation of both L1 and L2 orthography-to-phonology mappings during L2 written word recognition among low proficiency bilinguals.

<Insert Figure 1 about here>

This co-activation of both conversion systems during L2 written word recognition raises the question of spoken word recognition in L2, especially when lists of stimuli include cognate words, that is words sharing orthography but not systematically phonology between translation equivalents as mentioned above. French and English share the same alphabetic system, but their phonetic features slightly differ (Ryalls et al., 1995). Nevertheless, as shown by Marian, Spivey and Hirsch in 2003, low proficiency bilinguals tend to activate both languages in parallel, even in a monolingual situation of spoken word recognition. Thus, one might expect an interaction between modality and cognate status (see for cognate effect considering L2 words learned in oral modality Valente et al., 2018, see also for cognate effect in oral modality among high-proficiency bilinguals Wu et al., 2019). To summarise, there is now evidence that bilinguals simultaneously activate

lexical representations of both languages, both in written and oral modalities, with some interaction between modalities, notably through orthography-to-phonology mappings.

Critically, only a few bilingual models include both modalities. In most learning models as in expert-reader word recognition models, modalities are not considered. Shook and Marian proposed the Bilingual Language Interaction Network for Comprehension of Speech (BLINCS – Shook & Marian, 2013) to allow for intermodal activation. Their model is structured in phonological, phono-lexical, ortho-lexical and semantic levels, self-constructed by a self-adaptive algorithm. At each level, the representation of the two languages lies within the same network, which again makes it possible to account for competition between lexical representations of the two languages, particularly considering cognate words sharing orthographic representations in both languages. The ortho-lexical component of the model provides interaction between phonological and orthographic forms of spoken words and contains orthographic representations of lexical items. This makes it possible to consider the automatic activation of spelling during spoken word recognition, as well as the activation of phonological representations during reading.

Note that the impact of modality on word recognition has not been systematically explored *per se* either in L1 or in L2. In 2018, a mega-study (MEGALEX – Ferrand et al., 2018) was conducted in order to investigate visual and auditory French word recognition (L1). Participants performed visual and auditory lexical decision tasks with a very large number of items. While written words were recognized faster than spoken ones, accuracy was similar in both modalities. However, not all participants completed the tasks in both modalities on all items. To compare written and spoken inputs during the learning of new words in English as an L1, Nelson, Balass and Perfetti (2005) conducted an experiment of rare word learning. They found that participants required fewer trials to learn words in the orthographic condition than in the phonological one. Importantly, the accuracy scores testified to a modality effect in L1, with a more accurate

recognition of written words than spoken ones, especially when those words were learned in their orthographic form.

To our knowledge, only one study has examined the impact of modality in multilingual word processing. Veivo, Suomela-Salmi and Järvikivi (2015) studied a group of Finnish learners of French. Participants were bilingual speakers of Finnish and English, all expert readers, who learned French as a third language (L3) at school. One week after an online oral lexical decision task in French, participants performed a translation task from these same stimuli, always presented orally. Two months after these auditory tasks, the same translation task was performed, but with an orthographic presentation of the same stimuli. Given that the written modality was used more than the oral one in L3 learning at school, they thus highlighted a modality effect in L3, with higher accuracy scores for written words than spoken ones. Moreover, their results showed an interaction between L3 proficiency and modality, this modality effect being even greater for the low proficiency participants. The same group replicated those results in another study without a delay between modalities, using a repeated paradigm with a counterbalanced order of modalities. Note that the translation task necessitated semantic processing. The question then arises to what extent the first stage of word processing, namely word recognition, is sensitive to modality in intermediate proficiency bilinguals. More, it is important to evaluate whether modality and cognate effects interact during L2 word recognition among intermediate proficiency bilinguals, known to show a greater cognate effect in L2 (see for a meta-analytic review Lauro & Schwartz, 2017).

The aim of the current study was thus to examine modality and cognate effects – and their interaction, as well as the effect of L2 proficiency – in L2 word recognition across two experiments of lexical decision tasks in intermediate proficiency French-English bilinguals. There were no cognate words in the first experiment while the second one included cognate words. Given the predominance of written materials in L2 learning at school, our expectations were the followings. First, we expected a higher discrimination rate ( $d'$ ) for written words than spoken ones – this

parameter allowing us to distinguish possible individual bias towards YES or NO responses from the real participant's ability to differentiate words and pseudowords (Forrin et al., 2016; Hayes-Harb et al., 2010). Second, we anticipated a modality effect on L2 word recognition, with a more accurate and faster recognition of written words than spoken ones. Considering the narrow vocabulary breadth in intermediate proficiency bilinguals, we did not expect them to know exactly the same words, making it difficult to compare modalities in a between-group design. Therefore, we designed a paradigm using a repetition of the lists of stimuli, each participant performing the task twice, once in each modality. The order of modality was counterbalanced across participants, allowing us to exploratory compare the results across sessions. A session effect could thus indicate the benefit from one modality to the other, with higher accuracy scores and/or faster reaction times in the second session of word recognition. In addition, considering intermediate proficiency bilinguals, we expected a strong cognate effect on L2 written word recognition. Finally, based on the results obtained in L3 by Veivo et al. (2015), we expected those various effects to be dependent on L2 proficiency.

Our first experiment was designed to determine whether a modality effect (*i.e.*, written posited to be better than oral) and a session effect would be observed or not in L2, and if they interact with a proficiency effect (Experiment 1). Furthermore, we wanted to explore the links between modality and cognate status by checking whether the presence of cognate items modified the modality and session effects in English. Indeed, in French-English cognate words, orthographic overlap is much greater than phonological overlap (Experiment 2).

### **Experiment 1: Modality and session effects in L2**

#### ***Method.***

*Participants.* A total of 50 participants completed this experiment (37 females, 38 right-handed, mean age = 24 years, SD age = 4), all of them from several universities in the north of

France. All participants were native speakers of French and had learned English as an L2 at school in France. They had normal or corrected-to-normal vision and no hearing problems. None of them reported any kind of learning impairment. A participant with difficulties in the processing of written language, according to his reading-related background tests, was excluded from the analysis.

All participants gave their written informed consent and the study was approved by the Lille University Ethics Committee (Authorization # 2018 -263-S58).

The participants were divided randomly into two groups in order to counterbalance the order of presentation of modalities. We thus created a group of 24 participants first performing the task in the written modality and after in oral modality (W-O group). Then a group of 25 participants performed the task in the opposite order (O-W group).

*Background tests.* The following tests were administered to the participants: (1) an on-line questionnaire to describe their reading habits and experience with different languages; (2) a placement test in English to assess their level of proficiency in it, according to the Common European Framework of Reference (CEFR), using the Dialang test from Lancaster University (*Dialang*, s. d.)<sup>i</sup>; (3) reading-related and neuropsychological tests: reading and visual attention skills from the ECLA16+ battery (Gola-Asmussen et al., 2011), pseudoword reading from the EVALEC battery (Sprenger-Charolles et al., 2005) and non-verbal skills from the Raven Progressive Matrix test (Raven & Raven, 1998). Demographic data and results in the English placement tests, the reading-related and the neuropsychological tests are available in Table 1 (see Table S1 in supplementary materials for complete data). All statistics comparing the two groups regarding the tests were non-significant, except for laterality, with more left-handed participants in the W-O group ( $p = .04$ ).

<Insert Table 1 about here>

*Stimuli of Lexical Decision Task (LDT).* Words were selected from the SUBTLEX-UK database (Van Heuven et al., 2014), using the following selection criteria: frequency between 10 and 50 per million in oral form, 3-to-8 letter long, no homophones or homographs (in English but also in French), no plurals or conjugated forms. No cognate word was included. We then checked if those words were studied in a school context in France, our participants being French university students having learned English in a school context in France, using a database currently under development (ANR Grant-16-CE28-0009-01). This draft database contained all the words from 32 handbooks used during English lessons in French middle schools. From our first series of words from the SUBTLEX-UK database, we thus selected 44 words with frequency between 10 and 50 per million in written form in this latter database. Sixty-one percent were monosyllabic, others were disyllabic.

We used the software package Wuggy (Keuleers & Brysbaert, 2010) to create 44 pseudowords that were matched with the selected words. The following criteria were taken into account: no homophonic forms in either language, number of letters, phonemes and syllables, Levenshtein's average orthographic and phonological distances with the 20 closest neighbours, respectively OLD20 and PLD20 (calculated with LDCalc Software - Ferrand et al., 2010; Yarkoni, Balota, & Yap, 2008), frequencies and number of neighbours. Table 2a presents the pairings (see Table S2 in supplementary materials for complete pairings). Appendix 1 presents the complete lists of stimuli.

<Insert Table 2 about here>

The auditory stimuli were recorded by two native speakers (native English-speaking man and woman) using the software package Audacity (*Audacity - Copyright © 1989, 1991, s. d.*). Each stimulus had its own associated file and all the audio files were 44100-Hz stereo wav files and lasted about 1,000 ms. The male and female voices were counterbalanced across participants.

*Procedure.* All participants were tested in a quiet room at their university on the same testing apparatus. They performed two LDT in English, one in visual and one in auditory modality, one after the other, with 10 minutes break.

The experiment was run using the DMASTR software (DMDX version 5.1.5.3) developed at Monash University and at the University of Arizona by K.I. Forster and J.C. Forster (Forster & Forster, 2003).

For the visual LDT, stimuli were presented on a 15.6" Full HD laptop (Dell Precision Mobile 3520, Processor i5-7440HQ) with a refresh rate of 60 Hz and a resolution of 1,920 x 1,080 pixels. The monitor was placed at a distance of about 60 centimetres from the participants. Stimuli were presented in uppercase in Courier New (11-point font size). They appeared as black characters on a white background on the screen through a high-quality graphic card (NVIDIA Quadro M620).

For the auditory LDT, the apparatus was exactly the same as for the visual one, except that stimuli were played through speakers (Hercules XPS 2.030) with a high-quality audio soundcard included in the laptop motherboard (Dell Precision Mobile 3520).

The same list of stimuli was used in both modalities, with a counterbalanced order of modality presentation across participants. Thus, participants performed two sessions: one in each modality. There was a short break within each session (5 min). The stimuli were shuffled for each participant by the DMASTR software itself.

For the visual LDT, the sequence of each trial was the following: (1) a series of hashes (#####) appeared in the centre of the screen for 500 ms; (2) a stimulus (word or pseudoword) was presented in the centre of the screen until the participant's answer or for 4,000 ms maximum if no response was made. The inter-stimuli interval was 200 ms long.

For the auditory LDT, the procedure was identical to that used in the visual one, except that stimuli were played binaurally through speakers.

LDT were preceded by practice trials with verbal feedback from the examiner.

Participants responded on an XBOX 360 Controller, which does not have time delays with keyboards (Shimizu, 2002). We measured two dependent variables: response (correct or not) and reaction time (in milliseconds) for correct trials.

*Statistical analyses.* Data analysis was conducted with the software package R (R version 4.0.3) using different packages to perform descriptive and main analyses. We performed the following statistical analyses. First, we used the signal detection theory approach to analyse the discrimination rate between words and pseudowords for each participant (Green & Swets, 1966; Macmillan & Creelman, 1991). Discrimination rate ( $d'$ ) helped us to determine the extent to which each participant was able to discriminate the different stimuli. We analysed the  $d'$  of each participant depending on Modality, Session and L2Proficiency (corresponding to Dialang Level centered), using linear mixed-effect modelling (Baayen et al., 2008), with the function `lmer` of the package `lme4` (Bates et al., 2014) using the Akaike Information Criterion (AIC) to determine which model gave the best approximation of reality through a forward stepwise selection of fixed effects, starting with no fixed effect and adding them one by one<sup>ii</sup>, including the different possible two-way and three-way interactions. Concerning the random structure, it was simplified through a backward stepwise selection procedure, starting with the most complete random structure justified by the design, and stopping when all random effects resulting in nonconvergence were deleted (Barr et al., 2013). We presented in the result section the final chosen model only, with the different fixed effects (including all interactions involved) mentioned in the order in which we entered them in this model. To improve our confidence in the chosen model, we used Bayesian statistics, with the function `brm` of the `brms` package (Bürkner, 2017, 2018; Carpenter et al., 2017), to fit the same model and obtain 95% credible intervals (CrI) and posterior distributions for each estimate. We also used the `bayes_R2` function of the same package to calculate a Bayesian version of the  $R^2$  (Gelman et al., 2019).

Then, we analysed the Response (correct or incorrect, and thus the accuracy scores) of each participant on word trials only, to test if Modality and Session effects could be observed or not in L2 word recognition and if they depend on the level of L2Proficiency. We performed the main analyses on Response using generalized linear mixed-effect modelling, with the function `glmer` of the package `lme4` (Bates et al., 2014). We fitted a generalized linear mixed-effect model with a logit function to the response of each trial, using the AIC as above. We performed then the same type of Bayesian statistical analysis to obtain credible intervals and posterior distributions for each estimate.

Raw reaction times for correct word trials could not be analysed directly here. Indeed, the experimental design compared auditory and written lexical decisions, and reaction times are totally different in written and in oral modalities. Therefore, they are not directly comparable. In order to be able to examine the session effect, the reaction times were thus Z-transformed by modality. We have centred and reduced reaction times by considering each modality separately, to obtain Z-transformed reaction times (ZRT). With this transformation, positive ZRTs mean that words were recognized slower than the mean, and negative ZRTs mean that words were recognized faster than the mean, for the modality considered. We performed the main analyses on ZRT using linear mixed-effect modelling, with the function `lmer` of the package `lme4` (Bates et al., 2014). We fitted a linear mixed-effect model, using the AIC as above, with the same fixed-effect factors as above except for Modality. We performed then the same type of Bayesian statistical analysis as for  $d'$  to obtain credible intervals and posterior distributions for each estimate. When a session effect was highlighted, we then performed post-hoc analysis to compare the mean raw reaction times of our independent groups in each modality according to Session in order to underline an eventual interaction between Session and Modality.

### ***Results.***

*Discrimination rate ( $d'$ )*. The final linear mixed-effect model to the  $d'$  of each participant included, ~~with~~ Modality (written or oral), Session (1<sup>st</sup> or 2<sup>nd</sup>) and L2Proficiency (Dialang Level out of 1000) as fixed effects, and by-participants random intercepts (final best model according to the AIC:  $F(1,43) = 78.14$ ,  $p < .001$ ). Discrimination rates ( $d'$ ) were significantly lower in the oral modality (1.39,  $SD = .72$ ) than in written one (2.05,  $SD = .85$ ) and were also significantly lower in Session 1 (1.65,  $SD = .83$ ) than in Session 2 (1.79,  $SD = .88$ ). L2Proficiency effect was also significant, an increase of Dialang-Level score of 100 out of 1000 resulting in an increase of 0.35 of the  $d'$  of the participant.

The Bayesian analysis gave the same estimates for the parameters, with a tiny uncertainty, depicted by small estimate standard errors and thus relatively small width of credible intervals. This model seemed to correctly fit our data (bayes\_  $R^2 = .81$ ,  $SE = .04$ , 95% CrI = [0.72, 0.87]).

*Accuracy (response type)*. The final generalized linear mixed-effect model included ~~with~~ L2Proficiency and Modality (written or oral) as fixed effects, by-participants random intercepts and by-items random intercepts and random slope considering the Modality (final best model according to the AIC:  $AIC = 2986.9$ ,  $\chi^2(1) = 6.44$ ,  $p < .05$ ). This model highlighted a modality effect, accuracy scores being 3% lower in the oral modality (78%,  $SD = .41$ ) than in the written one (81%,  $SD = .39$ ). L2Proficiency effect was also significant: the more proficient, the more accurate.

The Bayesian analysis gave the same estimates for the parameters, with a relative tiny uncertainty, depicted by small estimate standard errors and thus relatively small width of credible intervals.

*Reaction time*. The final linear mixed-effect model included ~~with~~ Session (1<sup>st</sup> or 2<sup>nd</sup>) and L2Proficiency (Dialang Level out of 1000) as fixed effects, by-participants random intercepts and by-items random intercepts and random slope considering the Session (final best model according to the AIC:  $F(1,2826) = 25.48$ ,  $p < .001$ ). ZRT were significantly lower in Session 1 (.08,  $SD =$

1.05) than in Session 2 (-.08, SD = .94). L2Proficiency effect was also significant, an increase of Dialang-Level score of 100 out of 1000 resulting in a decrease of 0.15 of the ZRT.

The Bayesian analysis gave the same estimates for the parameters, with a tiny uncertainty, depicted by small estimate standard errors and thus relatively small width of credible intervals. The parameter estimates of the previous models ( $d'$ , Accuracy and ZRT), as well as the outputs from the Bayesian analyses as produced by these models are reported in Table 3.

<Insert Table 3 about here>

Given that reaction times were Z-transformed by Modality, considering both Sessions together, the mean ZRT was 0 and the standard deviation was 1 in each modality. Thus, it was not possible to include directly the two-way interaction between Modality and Session into the model. However, raw data suggested an interaction between Session and Modality. Therefore, we performed post-hoc analysis to compare the mean raw reaction times of our independent groups in each modality according to Session. They showed that there was a session effect with lower latencies in Session 2 than in Session 1 for the written modality ( $t(1368.3) = 6.812, p < .001$ ) but not for the oral one ( $t(1437.9) = -0.668, p = .504$ ).

## ***Discussion***

The aim of this first study was to determine whether a modality effect and a benefit between sessions can be observed in L2 and if they interact with a proficiency effect. We expected to evidence a modality effect, with a faster and more accurate recognition of written words than spoken ones in English. Moreover, we expected a proficiency effect and investigated its relation with other effects. Finally, we investigated the benefit from one modality to the other through a session effect.

We demonstrated a modality effect in all participants, with stronger discrimination and more accurate recognition of written words than spoken ones, whatever the order of presentation. This

suggests that the lexical knowledge of late learners of English as an L2 at school is closely linked to the modality of word presentation. There was a significant modality effect, as shown by a 3% easier recognition of words in written than in oral form, which seems however very consistent (cf. Bayesian analysis). Late learners of English as an L2 recognised ortho-lexical forms more easily than phono-lexical forms.

Moreover, we failed to find a session effect, which would have reflected a benefit from one modality to the other in terms of accuracy. These results suggest that L2 learners use their reading skills to create strong L2 ortho-lexical representations from the written input. These representations seem more robust (*cf.* modality effect) than phono-lexical ones, which is consistent with the imprecision of the phono-lexical representations found in late L2 learners (Cook et al., 2016 ; Darcy et al., 2013). It is also congruent with the results of Veivo and colleagues (2015). They found in Finnish learners of French as an L3 completing a translation task that imprecise phono-lexical representations were sometimes associated with accurate ortho-lexical representations.

However, in terms of latencies, we found a session effect only in the written modality. This effect denoted that written words were recognized faster in Session 2, thus when spoken words were recognised first. This indicates a benefit from oral to written modality in terms of reaction times. Thus, hearing an English word activates a lexical representation which intermediate proficiency bilinguals use for written word recognition. Thus, when they see the same word just after hearing it, they recognize it faster, the lexical representation being pre-activated. On the contrary, their lexical representations of English words encountered in written form do not accelerate their auditory recognition. A possible explanation is that several orthography-to-phonology mappings can be applied for these English written words. It is likely that the phonological product of these mappings does not match with phono-lexical representations. Thus, when they hear the same words just after, they derive no benefit from the first presentation. Overall,

our data suggest that ortho-lexical representations are not tightly connected to phono-lexical ones in English as an L2 among French-English intermediate proficiency bilinguals.

Finally, we found a proficiency effect: the highest L2 proficiency was associated to the highest discrimination rates and accuracies, and the fastest latencies. Critically, the interactions between L2Proficiency and Modality on the one hand, and between L2Proficiency and Session on the other hand, were non-significant, modality and session effects existing whatever the L2 proficiency.

Given this modality effect in L2, and especially in view of the absence of session effect in terms of accuracy despite a benefit from oral to written modality in terms of reaction times in this first experiment, the question arose regarding the influence of cognate words on L2 word recognition in both modalities. Indeed, the overlap between two languages sharing the same alphabetic system is more often based on spelling than on phonology, as is the case between French and English. Thus, cognate words in these two languages are closer orthographically than phonologically. However, the cognate effect has been studied mostly in the written modality and, crucially, no comparison of this effect according to modality has been performed yet. In the Experiment 2, we analysed the links between modality and cognate effects by checking whether the presence of cognate items altered the modality effect in English, as well as determining whether it modified the benefit between sessions or whether it interacted with L2 proficiency.

## **Experiment 2: Links between modality and cognate status in L2**

### ***Method.***

*Participants.* The same participants as in Experiment 1 performed Experiment 2, about 15 minutes after the first one. All participants performed Experiment 2 with cognate words included into the lists of stimuli after Experiment 1 without cognate words. We decided not to counterbalance the order of experiments to avoid an effect of cognate words during the experiment

without cognate words – linked with a previous activation due to the experiment including cognate words.

*Stimuli of LDT.* New words were selected from the SUBTLEX-UK database (Van Heuven et al., 2014), excluding all the non-cognate words used in the first experiment and using the following selection criteria: frequency between 10 and 50 per million in oral form, 3-to-12 letter long, no homophones or homographs, no plurals or conjugated forms. We then checked if those words were studied in a school context in France with the same handbook database as in Experiment 1. From our first series of words from the SUBTLEX-UK database, we thus selected 60 words with frequencies between 10 and 50 per million in written form according to this latter database. Thirty-five percent were monosyllabic, 43% were disyllabic, 15% were trisyllabic, and the others were quadri-syllabic<sup>iii</sup>. Word frequencies were identical between Experiments 1 and 2 ( $t(101.95) = -0.428, p > .20$ ).

Among these 60 words, 30 were cognate words (with a Levenshtein distance lower than 3 between them and their translation equivalents) and 30 were non-cognate words (with a Levenshtein distance greater than 3 between them and their translation equivalents).

In order to evaluate the effect of the amount of orthographic overlap between translation equivalents, we selected among the 30 cognate words 15 identical cognate words (with exactly the same orthographic form in both languages, and thus a Levenshtein distance of 0 between them; e.g. “rare”) and 15 non-identical cognate words (with slight differences in their orthographic forms between the two languages, and thus a Levenshtein distance lower than 3 between them; e.g.: “honnête” in French / “honest” in English).

We used the same procedure as in Experiment 1 to create 60 pseudowords, strictly matched with those words, using exactly the same parameters. Table 2b presents the pairings (see Table S3

in supplementary materials for complete pairings). Appendix 2 presents the complete lists of stimuli.

*Procedure.* The procedure used was identical to that in Experiment 1.

*Statistical analyses.* The analyses performed were the same as in Experiment 1, except that we performed the analyses including Cognateness (Cognate or Non cognate words) as fixed-effect in each analysis, in addition with fixed-effects of Experiment 1, as well as the interaction between all those factors. Secondly, we performed additional analyses including only cognate words, in order to determine more precisely the influence of Cognate-type (Identical or Non-identical cognate words), and thus to evaluate the effect of the amount of orthographic overlap between translation equivalents.

### ***Results.***

*Discrimination rate ( $d'$ ).* The final linear mixed-effect model to the  $d'$  of each participant included Modality (written or oral), Cognateness (cognate or non-cognate words), L2Proficiency and the interaction between Modality and Cognateness as fixed effects, and by-participants random intercepts (final best model according to the AIC:  $F(1,44) = 33.82$ ,  $p < .001$ ). Discrimination rates ( $d'$ ) were significantly lower in the oral modality (1.69,  $SD = .75$ ) than in the written one (2.65,  $SD = .84$ ) and significantly lower for non-cognate words (2.02,  $SD = .88$ ) than for cognate ones (2.32,  $SD = .96$ ). L2Proficiency effect was also significant, an increase of Dialang-Level score of 100 out of 1000 resulting in an increase of 0.27 of the  $d'$  of the participant. Interestingly, the interaction between Modality and Cognateness was significant,  $d'$  being significantly higher for cognate words than non-cognate ones in written modality ( $t(92) = 2.53$ ,  $p < .01$ ), but not in oral modality ( $t(92) = .84$ ,  $p = .406$ ).

The Bayesian analysis gave the same estimates for the parameters, with a tiny uncertainty, depicted by small estimate standard errors and thus relatively small width of credible intervals. This model seemed to correctly fit our data (bayes\_  $R^2 = .72$ ,  $SE = .03$ , 95% CrI = [0.65, 0.77]).

*Accuracy (response type).* The final generalized linear mixed-effect model included L2Proficiency, Modality (written or oral), Cognateness (cognate or non-cognate words) and the interaction between Modality and L2 Proficiency as fixed effects, and by-participants random intercepts<sup>iv</sup> (final best model according to the AIC:  $AIC = 3876.9$ ,  $\chi^2(1) = 9.96$ ,  $p < .01$ ). This model evidenced a modality effect, accuracy being lower in the oral modality (83%,  $SD = .38$ ) than in the written one (87%,  $SD = .33$ ). L2-Proficiency effect was also significant: the more proficient, the more accurate. Interestingly, the interaction between L2-Proficiency and Modality was significant, the more a participant was proficient in L2, the higher he/she was accurate, notably in oral modality (see Figure 2). This model highlighted also a cognate effect, accuracy scores being lower for non-cognate words (83%,  $SD = .38$ ) than for cognate ones (87%,  $SD = .33$ ).

<Insert Figure 2 about here>

The Bayesian analysis gave the same estimates for the parameters, with a relative tiny uncertainty, depicted by small estimate standard errors and thus relatively small width of credible intervals.

*Reaction time.* The final linear mixed-effect model included Session (1<sup>st</sup> or 2<sup>nd</sup>), L2Proficiency (Dialang Level out of 1000) and their interaction as fixed effects, by-participants and by-items random intercepts (final best model according to the AIC:  $F(1,4080) = 4.90$ ,  $p < .05$ ). ZRT were significantly lower in Session 1 (.11,  $SD = 1.09$ ) than in Session 2 (-.11,  $SD = .89$ ). L2Proficiency effect was also significant, an increase of Dialang-Level score of 100 out of 1000 resulting in a decrease of 0.13 of the ZRT.

The Bayesian analysis gave the same estimates for the parameters, with a tiny uncertainty, depicted by small estimate standard errors and thus relatively small width of credible intervals<sup>v</sup>. The parameter estimates of the previous models ( $d'$ , Accuracy and ZRT), as well as the outputs from the Bayesian analyses as produced by these models are reported in Table 4.

<Insert Table 4 about here>

In the same way as for Experiment 1, it was not possible to include the two-way interaction between Modality and Session directly into the model. Thus, we performed post-hoc analysis to compare the mean raw reaction times of our independent groups in each modality according to Session. They showed that there was a session effect with lower latencies in Session 2 than in Session 1 for the written modality ( $t(1992.9) = 8.392, p < .001$ ) but not for the oral one ( $t(2033) = 1.425, p = .154$ ).

Likewise, raw data of reaction times suggested an interaction between Modality and Cognateness, which could explain the absence of cognate effect observed in the model. But, this interaction could not be included directly into the model, due to the Z-transformation by-modality of reaction times. Thus, we performed also post-hoc analysis to compare the mean raw reaction times of our independent groups in each modality according to Cognateness. They showed that there was a cognate effect, with cognate words being faster recognized than non-cognate words in the written modality ( $t(2141.2) = -1.993, p < .05$ ) but not in the oral one ( $t(2017.5) = -0.490, p = .624$ ).

*Further analysis (cognate type of word influence).* Because the overlap between French and English is greater in orthography than in phonology, we further explored this cognate effect on word recognition in both modalities by comparing identical and non-identical cognate words. Thus, we performed the same analyses as previously (for accuracy and ZRT) but including only cognate words. To Response, the final generalized linear mixed-effect model included L2Proficiency,

Modality (written or oral), Cognate-type (identical or non-identical cognate words), and the interactions between L2Proficiency and Modality -type and between Modality and Cognate as fixed effects, by-participants random intercepts<sup>iii</sup> (final best model according to the AIC: AIC =1753.3,  $\chi^2(1) = 6.80$ ,  $p < .01$ ). This model revealed a modality effect with written words (89%, SD = .31) being better recognized than spoken ones (85%, SD = .36). L2Proficiency effect was also significant: the more proficient, the more accurate. Interestingly, the interaction between L2Proficiency and Modality was significant, the more a participant was proficient in L2, the higher he/she was accurate, notably in oral modality (see Figure 3). This model highlighted also a cognate-type effect, accuracy being lower for non-identical cognate words (85%, SD = .35) than for identical cognate ones (89%, SD = .31). Interestingly, the interaction between Modality and Cognate-type was also significant, identical cognate words being better recognized than non-identical ones in written modality, but not in oral one.

<Insert Figure 3 about here>

The Bayesian analysis gave the same estimates for the parameters, with a relative tiny uncertainty, depicted by small estimate standard errors and thus relatively small width of credible intervals.

Finally, to ZRT, the final linear mixed-effect model included Session (1<sup>st</sup> or 2<sup>nd</sup>) and L2Proficiency (Dialang Level out of 1000) as fixed effects, by-participants and by-items random intercepts (final best model according to the AIC:  $F(1,2095) = 23.28$ ,  $p < .001$ ). ZRT were significantly lower in Session 1 (.10, SD = 1.10) than in Session 2 (-.09, SD = .88). L2Proficiency effect was also significant, an increase of Dialang-Level score of 100 out of 1000 resulting in a decrease of 0.09 of the ZRT.

The Bayesian analysis gave the same estimates for the parameters, with a tiny uncertainty, depicted by small estimate standard errors and thus relatively small width of credible intervals. The

parameter estimates of the previous models (Accuracy and ZRT), as well as the outputs from the Bayesian analyses as produced by these models are reported in Table 5.

<Insert Table 5 about here>

As in the previous analysis, we performed post-hoc analysis to compare the mean raw reaction times of our independent groups in each modality according to Session. They showed that there was a session effect with lower latencies in Session 2 than in Session 1 for the written modality ( $t(994.87) = 5.894, p < .001$ ) but not for the oral one ( $t(1001) = .392, p = .695$ ).

For the same reason as previously, we performed also post-hoc analysis to compare the mean raw reaction times of our independent groups in each modality according to Cognate-type. They showed that there was no cognate-type effect ( $t(2119.7) = .595, p = .552$ ), although identical cognate words were recognized faster than non-identical cognate words in the written modality ( $t(1079.1) = 3.239, p < .01$ ), and slower in the oral one ( $t(1036.7) = -2.481, p < .05$ ).

### ***Discussion***

The aim of this second experiment was to explore the links between modality and cognate status by checking whether the presence of cognate items modified the modality effect and benefit between modalities in English, and whether it interacted with L2 proficiency.

Importantly, while the lists of stimuli included both non-cognate and cognate words, this experiment revealed a comparable modality effect as in Experiment 1 (without cognates), *i.e.*, a more accurate and faster recognition of written words than spoken ones. Furthermore, the mean accuracy scores exhibited a better recognition of words when the lists included cognate words. Since the two experiments were conducted approximately 15 minutes apart, with the same participants, this gain in accuracy could, of course, be related to a practice effect. However, since the lists of stimuli were completely distinct from one experiment to the other, without any overlap,

this effect can only be limited<sup>vi</sup> and would rather be reduced to a training or habituation to the task effect, and in particular to listening to words in L2 - in this case English.

Overall, cognate words elicited shorter response latencies and fewer errors than non-cognate words. This advantage of cognate words has already been studied in written modality (see for a meta-analytic review Lauro & Schwartz, 2017). This effect may be explained by the fact that orthographic and semantic overlaps match for cognate words.

To our knowledge, the current study is the first to compare cognate word recognition across modalities. We obtained higher accuracy scores for cognate words than non-cognate ones whatever the modality, but lower latencies and higher discrimination rates for cognate words only in written modality. This suggests an interference in oral modality, reflecting activation of the ortho-lexical representation of spoken words during their recognition. Given the incongruence between French and English orthography-to-phonology mappings and thus the absence of a phonological overlap between them, cognate words do not seem to accelerate word recognition as easily in the oral modality as in the written one. Moreover, cognate words seemed to help the discrimination between words and pseudowords only in written modality, which tends to indicate that the orthographic overlap between translation equivalents of cognate words make this discrimination easier. Thus, the well-known facilitation effect of cognate words seems to exist only in the written modality in terms of discrimination, reaction time and accuracy. In the oral modality, cognate effect seems to exist only in terms of accuracy, cognate words hampering, or at least not accelerating, spoken word recognition.

Critically, the current experiment highlighted an interaction between Modality and L2 Proficiency: the lower the L2 Proficiency, the greater the modality effect – *i.e.*, the greater the difference in accuracy between modalities (see Figure 2), interaction also present when considering only cognate words (see Figure 3). Thus, there were: a) a modality effect only for lower proficiency

participants; b) a cognate facilitation effect only in the written modality; and c) a modality effect limited to lower proficiency participants and cognate items. This suggests that our population of late learners of L2 in written form is particularly sensitive to the cognate effect. As in the first experiment, there was a session effect in the written modality in terms of reaction time, with written words being recognized faster when spoken word recognition was performed first. This indicates a benefit from oral to written modality in terms of reaction time. This effect could be due to the two characteristics of English language we highlighted previously. Firstly, it could be linked with the incongruence of orthography-to-phonology mappings between French and English, hampering the activation of correct phono-lexical representation from written words, and thus preventing the faster recognition of spoken words when written word recognition was performed first, due to competition between phonological forms. Secondly, this could be a consequence of the inconsistency of English language, combined with multilingualism. Indeed, knowing different languages, bilinguals can create several phonological representations from the same written input: a) because of the incongruence between French and English orthography-to-phonology mappings (see Figure 1 top and bottom arrows), and b) because the inconsistency of orthographies can lead to the creation of different phonological representations from the same written input using diverse orthography-to-phonology possible mappings of one language (see Figure 1 middle and bottom arrows for English example). Thus, the created phonological representations of written words might be inaccurate and cannot help them in the later spoken word recognition. On the contrary, from an oral input, they have several cues to determine the language to consider. Therefore, they could create precise orthographic representations of spoken words, these pre-activated representations allowing a faster recognition of the written words presented afterwards. In the same time, the benefit from oral to written modality seems to indicate that visual word recognition involves English phonological skills. The result of the orthography-to-phonology mappings of someone who only knows English will probably be closer to correct form than that of a French speaker learning English (and who

could use both orthography-to-phonology mappings to create a phonological output from written English words).

In addition, this experiment showed that accuracy is higher for identical cognate words than non-identical ones, in the written modality only. Moreover, identical cognate words were recognized faster than non-identical ones in the written modality, whereas they were recognized more slowly in the oral one. This suggests that cognate status has an impact on the modality effect, which depends on the quantity of orthographic overlap (identical cognate words being identical in written modality only). Nevertheless, a simple partial overlap is enough to reinforce the activation of ortho-lexical representations, making the recognition of written words easier than that of spoken ones. In terms of reaction time, a complete orthographic overlap allows a faster recognition of written words, whereas it hampers spoken word recognition, probably due to the incongruence of both orthographies, thus leading to a competition between phonological forms of the two languages.

Finally, as in the experiment without cognate words, there was a session effect in terms of reaction time only in the written modality, with written words being recognized faster when spoken word recognition was performed first. This suggests that orthographic inconsistency in English from written to oral modality made it difficult for our participants to create correct phonological representations from written words.

### ***General discussion and conclusion***

The main aim of the present study was to examine the impact of modality on L2 word recognition – and its interaction with cognate effect – across lexical decision tasks in intermediate proficiency French-English bilinguals. To summarize, this study highlighted, whatever the presence or not of cognate items: (a) modality and L2 Proficiency effects on word-pseudoword discrimination and on word-recognition accuracy and latencies; and (b) a session effect on word recognition latencies only in the written modality.

Therefore, as expected, we found a modality effect in L2, with written words being recognized more accurately than spoken ones, irrespective of the presence or not of cognate items in the lists of stimuli. This effect is probably linked with different activation mechanisms according to the modality considered – lexical in written modality and sub-lexical in oral one. This modality effect is linked with L2 proficiency and cognate status – and their interaction (see Figures 2 and 3). These results are congruent with the results of Veivo et al. in 2015 who found a modality effect in L3, with the same pattern of results, and an interaction between L3 proficiency and modality. Critically, the current study highlighted an interaction between Modality and L2 Proficiency in Experiment 2 only (considering all items as well as cognate items only): the lower the L2 Proficiency, the greater the modality effect – *i.e.*, the greater the difference in accuracy between modalities. This interaction was not found in Experiment 1, which did not contain cognate items. This seems to indicate that the presence of cognate items modified the link between modality and proficiency effects and is, at least partially, responsible of this greater modality effect among lower proficiency participants.

Considering the specificity of cognate words, our findings reveal the well-known cognate facilitation effect in terms of accuracy and reaction time, which is mostly dependent on the amount of orthographic overlap between translation equivalents of both languages.

In terms of reaction time, we found a session effect, indicating a benefit from oral to written modality, *i.e.*, written words were recognised faster when they had been presented orally first. The results suggest that the activation of the ortho-lexical representations of words is greater than that of phono-lexical ones. This also might be due to the orthographic overlap between translation equivalents of the two languages. Note that a complete orthographic overlap allows the faster recognition of written words, whereas it hampers spoken word recognition. Those findings suggest competition between phonological forms of the two languages, due to their incongruence. Indeed, the BIA-d model seems to be modal-dependent.

Critically, this first experiment comparing cognate word recognition across modalities suggest that the well-known cognate facilitation effect should be put into perspective. While it exists in the written modality, it is reversed in oral modality, since the presence of a complete orthographic overlap between translation equivalents of the two languages hampers spoken word recognition. This is the first evidence of a two-sided (different in both modalities) cognate effect that needs further investigation.

The question then arises about the existence of modal-dependent or amodal lexical representation of L2 words, as this debate has not yet been decided in L1 (see for L1 French: (Sauval et al., 2018); see for L1 English: (Balota et al., 1999)). Indeed, in L2, especially among intermediate proficiency bilinguals, the access to the lexical representation of new words does not seem to be effective without taking modality into account, resulting in a modality effect in favour of the written one. Moreover, due to the practicalities of L2 learning at school, there seems to be a bias in exposure to the written modality. Thus, ortho-lexical representations seem to be more easily and specifically activated than phono-lexical ones.

There are some limitations to our study. First, because different language pairs are characterized by both different inconsistency and incongruence levels, we could expect those results to be modulated by the language pair considered. Second, our experimental design didn't allow a complete analysis of the modality effect on latencies. Yet, an analysis of the difference in latencies would certainly be more precise. To this end, specific experiments should be designed as a direct comparison between modalities is not possible, due to processing latency differences.

In conclusion, due to the impact of modality on word recognition skills, our results plaid in favour of modal-dependent lexical representations of English as an L2 among French-English late bilinguals. Critically, our results indicate that the cognate facilitation effect needs to be nuanced. Indeed, the cognate effect does not always seem to lead to facilitation. On the contrary, it is a

modal-dependent effect: facilitating in the written modality but hampering spoken word recognition. Such results open new perspectives for research into how words learned through reading are recognized when spoken. Future research should take various parameters into account: modality, of course, but also cognate status and cognate-type, *i.e.*, the degree of orthographic overlap between translation equivalents of both languages. Moreover, does this modality effect have any impact on specific populations who have difficulties in written string processing, such as dyslexic readers? There is also the issue of the impact of L2 proficiency on word recognition in both modalities. Research is needed on low proficiency bilinguals, such as beginner learners of English as an L2.

### **Supplementary material.**

Table S1. Complete demographic data and background tests: comparison of groups.

Table S2. Complete pairing parameters for stimuli in Experiment 1.

Table S3. Complete pairing parameters for stimuli in Experiment 2.

For supplementary material accompanying this paper, visit

[www.cambridge.org/core/journals/bilingualism-language-and-cognition](http://www.cambridge.org/core/journals/bilingualism-language-and-cognition)

### **References.**

*Audacity*—Copyright © 1989, 1991. (s. d.).

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(4), 390-412. <https://doi.org/10.1016/j.jml.2007.12.005>

Balota, Watson, J. M., Duchek, J. M., & Ferraro, F. R. (1999). Cross-modal semantic and homograph priming in healthy young, healthy old, and in Alzheimer's disease

- individuals. *Journal of the International Neuropsychological Society*, 5(7), 626-640.  
<https://doi.org/10.1017/S1355617799577060>
- 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). <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2014). *Fitting linear mixed-effects models using lme4*. arXiv preprint arXiv:1406.5823. <https://arxiv.org/pdf/1406.5823.pdf>
- Bürkner, P.-C. (2017). brms : An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), 1-28. <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.
- Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan : A Probabilistic Programming Language. *Journal of Statistical Software*, 76(1), 1-32. <https://doi.org/10.18637/jss.v076.i01>
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC : A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204-256. <https://doi.org/10.1037/0033-295X.108.1.204>
- Commissaire, E., Duncan, L. G., & Casalis, S. (2019). Investigating pseudohomophone interference effects in young second-language learners. *Journal of Experimental Child Psychology*, 180, 1-18. <https://doi.org/10.1016/j.jecp.2018.11.010>
- Cook, S. V., Pandza, N. B., Lancaster, A. K., & Gor, K. (2016). Frontiers | Fuzzy Nonnative Phonolexical Representations Lead to Fuzzy Form-to-Meaning Mappings | Psychology. *Frontiers in psychology*, 7, 1345.

- Darcy, I., Daidone, D., & Kojima, C. (2013). Asymmetric lexical access and fuzzy lexical representations in second language learners | John Benjamins. *The Mental Lexicon*, 8(3), 372-420.
- Dialang*. (s. d.). Consulté 10 décembre 2019, à l'adresse <https://dialangweb.lancaster.ac.uk/>
- Dijkstra, T., Van Heuven, W., & Grainger, J. (1998). Simulating cross-language competition with the bilingual interactive activation model. *Psychologica Belgica*, 38(3-4), 177-196.
- 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(3), 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. (2019). Multilink : A computational model for bilingual word recognition and word translation | Bilingualism : Language and Cognition | Cambridge Core. *Bilingualism: Language and Cognition*, 22(4), 657-679.
- Ferrand, L., Méot, A., Spinelli, E., New, B., Pallier, C., Bonin, P., Dufau, S., Mathôt, S., & Grainger, J. (2018). MEGALEX : A megastudy of visual and auditory word recognition. *Behavior Research Methods*, 50(3), 1285-1307. <https://doi.org/10.3758/s13428-017-0943-1>
- Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., Augustinova, M., & Pallier, C. (2010). The French Lexicon Project : Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods*, 42(2), 488-496. <https://doi.org/10.3758/BRM.42.2.488>
- Forrin, N. D., Groot, B., & MacLeod, C. M. (2016). The d-Prime directive : Assessing costs and benefits in recognition by dissociating mixed-list false alarm rates. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(7), 1090-1111. <https://doi.org/10.1037/xlm0000214>

- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior research methods, instruments & computers*, 35(1), 116-124.
- 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>
- Gola-Asmussen, C., Lequette, C., Pouget, G., Rouyer, C., & Zorman, M. (2011). *ECLA-16+ : Évaluation des compétences de lecture chez l'adulte de plus de 16 ans*.  
[https://scholar.google.fr/scholar?hl=fr&as\\_sdt=0%2C5&q=Gola-Asmussen%2C+Lequette%2C+Pouget%2C+Rouyer%2C+%26+Zorman%2C+2011&btnG=](https://scholar.google.fr/scholar?hl=fr&as_sdt=0%2C5&q=Gola-Asmussen%2C+Lequette%2C+Pouget%2C+Rouyer%2C+%26+Zorman%2C+2011&btnG=)
- 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 (Éds.), *Language Acquisition and Language Disorders* (Vol. 52, p. 267-283). John Benjamins Publishing Company. <https://doi.org/10.1075/lald.52.18gra>
- Green, D. M., & Swets, J. A. (1966). *SDT and Psychophysics.pdf*.  
[http://andrei.gorea.free.fr/Teaching\\_fichiers/SDT%20and%20Psychophysics.pdf](http://andrei.gorea.free.fr/Teaching_fichiers/SDT%20and%20Psychophysics.pdf)
- Hayes-Harb, R., Nicol, J., & Barker, J. (2010). Learning the Phonological Forms of New Words : Effects of Orthographic and Auditory Input. *Language and Speech*, 53(3), 367-381. <https://doi.org/10.1177/0023830910371460>
- Keuleers, E., & Brysbaert, M. (2010). Wuggy : A multilingual pseudoword generator. *Behavior Research Methods*, 42(3), 627-633. <https://doi.org/10.3758/BRM.42.3.627>
- Lauro, J., & Schwartz, A. I. (2017). Bilingual non-selective lexical access in sentence contexts : A meta-analytic review—ScienceDirect. *Journal of Memory and Language*, 92, 217-233.
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory : A user's guide*.

- Marian, V., Spivey, M., & Hirsch, J. (2003). Shared and separate systems in bilingual language processing : Converging evidence from eyetracking and brain imaging. *Brain and Language*, 86(1), 70-82. [https://doi.org/10.1016/S0093-934X\(02\)00535-7](https://doi.org/10.1016/S0093-934X(02)00535-7)
- Nelson, J. R., Balass, M., & Perfetti, C. A. (2005). Differences between written and spoken input in learning new words. *Written Language & Literacy*, 8(2), 25-44. <https://doi.org/10.1075/wll.8.2.04nel>
- Pivneva, I., Mercier, J., & Titone, D. (2014). Executive control modulates cross-language lexical activation during L2 reading : Evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40(3), 787-796. <https://doi.org/10.1037/a0035583>
- Raven, J., & Raven, J. C. (1998). *Manual for Raven's progressive matrices and vocabulary scales*.
- Ryalls, J., Provost, H., & Arsenault, N. (1995). Voice Onset Time production in French-speaking aphasics—ScienceDirect. *Journal of Communication Disorders*, 28(3), 205-215.
- Sauval, K., Perre, L., & Casalis, S. (2018). Phonemic feature involvement in lexical access in grades 3 and 5 : Evidence from visual and auditory lexical decision tasks. *Acta Psychologica*, 182, 212-219. <https://doi.org/10.1016/j.actpsy.2017.12.002>
- Shimizu, H. (2002). Measuring keyboard response delays by comparing keyboard and joystick inputs. *Behavior research methods, instruments & computers*, 34(2), 250-256.
- Shook, A., & Marian, V. (2013). The Bilingual Language Interaction Network for Comprehension of Speech. *Bilingualism: Language and Cognition*, 16(2), 304-324. <https://doi.org/10.1017/S1366728912000466>

- Sprenger-Charolles, L., Colé, P., Béchennec, D., & Kipffer-Piquard, A. (2005). French normative data on reading and related skills : From 7 to 10 year-olds. *European Review of Applied Psychology*, 55, 157-186.
- Valente, D., Ferré, P., Soares, A., Rato, A., & Comesaña, M. (2018). Does phonological overlap of cognate words modulate cognate acquisition and processing in developing and skilled readers? *Language Acquisition*, 25(4), 438-453. <https://doi.org/10.1080/10489223.2017.1395029>
- Van Heuven, W. J. B., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English : The Quarterly Journal of Experimental Psychology : Vol 67, No 6. *The Quarterly Journal of Experimental Psychology*, 67(6), 1176-1190.
- Veivo, O., Suomela-Salmi, E., & Järvikivi, J. (2015). Orthographic bias in L3 lexical knowledge : Learner-related and lexical factors | John Benjamins. *Language, Interaction and Acquisition*, 6(2), 270-293.
- Wu, J., Chen, Y., van Heuven, V. J., & Schiller, N. O. (2019). Dynamic effect of tonal similarity in bilingual auditory lexical processing. *Language, Cognition and Neuroscience*, 34(5), 580-598. <https://doi.org/10.1080/23273798.2018.1550206>
- Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N : A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, 15(5), 971-979. <https://doi.org/10.3758/PBR.15.5.971>
- Yudes, C., Macizo, P., & Bajo, T. (2010). Cognate effects in bilingual language comprehension tasks : NeuroReport. *NeuroReport*, 21(7), 507-512.
- Ziegler, J. C., & Goswami, U. (2005). Reading Acquisition, Developmental Dyslexia, and Skilled Reading Across Languages : A Psycholinguistic Grain Size Theory. *Psychological Bulletin*, 131(1), 3-29. <https://doi.org/10.1037/0033-2909.131.1.3>

## Appendix 1. Stimuli in Experiment 1.

### **Words.**

aim, anger, attic, baker, bean, belt, blind, breath, bunch, ceiling, chicken, crew, dish, dry, dull, duty, frame, garlic, gift, glad, guilty, heaven, honey, hook, ladder, leaf, level, loss, mistake, mood, neck, purple, sand, shame, sharp, shoulder, sight, sink, smoke, truth, wet, wind, wing, wool.

### **Pseudowords.**

arker, attay, aze, bealing, beft, blith, bozer, brooth, bry, burge, chacken, cred, dimp, duse, frad, frane, geardy, goft, hool, hotey, hounen, kond, lammer, lodel, loff, mistyle, murgle, musy, nell, rif, rin, shail, sharf, shielder, sint, sitch, smike, tink, trith, wartic, wess, wike, wook, yond.

## Appendix 2. Stimuli in Experiment 2.

### Words.

French-English identical cognates: accident, application, architecture, excuse, garage, global, incident, menu, rare, regret, rival, section, signal, signature, tropical.

French-English non-identical cognates: access, adult, alcohol, apartment, classical, economy, exchange, flame, honest, majority, onion, paradise, powder, private, sense.

French-English non-cognates: brand, cow, currency, deep, doll, drum, drunk, faith, handsome, joke, lamb, lift, lorry, luck, nasty, noisy, pride, rogue, roof, seaside, shape, shed, smelly, spoon, steam, tiny, towel, wealth, wicked, witness.

### Pseudowords.

abbliration, accibate, afress, alpowel, anilement, arthitacture, blunk, brape, brate, clammical, currenby, drom, ecanory, exbuse, exchaine, fearn, feep, flape, fow, gacked, garock, glibal, grubical, handsall, homel, horest, incigate, jore, lage, lide, lidy, lunk, moof, nispy, odult, oroon, paladent, poharity, ponu, prinate, purder, rebrel, ricel, roise, rore, sanse, seasile, seclion, sheed, signanise, sildal, smanty, spoop, stean, tride, turry, wayness, weafed, wisty, woll.

---

<sup>i</sup> Dialang is an online test proposing the evaluation of 5 communication skills: oral comprehension (*i.e.*, comprehension skill from oral sentences), written expression (*i.e.*, capacity to complete sentences with the correct word, to decide which sentence is not a part from a text, to judge if sentences are correct or not), written comprehension (*i.e.*, comprehension skill from written sentences), grammatical structures (*i.e.*, grammatical skill) and vocabulary (*i.e.*, capacity to choose the correct word to express an idea, a specific meaning, in written form). Those skills are evaluated, resulting in a level according to the Common European Framework of Reference (CEFR). Dialang propose also a self-assessment, to compare the self-estimation with the results of the tests and a first level-assessment, with a lexical decision task, leading to a score out of 1000. In this study, our participants performed the self-assessment, the first level-assessment (called Dialang\_Level here), the oral comprehension test and the vocabulary test, all of the results being highly correlated (see Table 1 and Table S1). The validity of this test was analysed through a comparison with the LEXTALE test among 49 participants. There scores were correlated at .68.

<sup>ii</sup> We conducted also the same analysis with a backward procedure, starting with the full model. We obtained exactly the same results in each analysis.

<sup>iii</sup> According to the difficulty to select cognate words in this specific range of frequency, we needed to extend our selection criteria with words of 1 to 4 syllables. This would probably have an influence on raw latencies, but no influence on our results, the same words being used in both modalities.

<sup>iv</sup> We didn't include by-item random intercepts in the model because our backward stepwise selection procedure showed us that the variance on the items was too low ( $< .002$ ). Thus, the inclusion of by-item random intercepts prevented the model from convergence.

- 
- v Note that the mixed-effect model showed a significant interaction between Session and L2Proficiency, but with a parameter for which the credible interval contained 0.00.
- vi Note that the pattern of results was similar in both experiments when we considered only the subset of non-cognate words in the second experiment, consisting in a replication of our findings in Experiment 1.