Psychonomic Bulletin & Review (2019) 26:353–359
https://doi.org/10.3758/s13423-018-1518-6

BRIEF REPORT

Disentangling cross-language orthographic neighborhood from markedness effects in L2 visual word recognition

Eva Commissaire[1] · Jean Audusseau[1] · Séverine Casalis[2]

Published online: 5 September 2018

Abstract

Previous research has reported that lexical access in bilinguals is language non-selective. In the present study, we explored the extent to which cross-language orthographic neighborhood size (N-size) effects, an index of language non-selectivity, should be dissociated from markedness effects, a sub-lexical orthographic variable referring to the degree of language- shared (unmarked) versus specific (marked) orthography. Two proficiency groups of French/English bilinguals performed an English (L2) lexical decision task with three word and non-word conditions: (1) English words with large French N-size/unmarked orthography (*price*), (2) small French N-size/unmarked orthography (*drive*), and (3) small French N-size/marked orthography (*write*). Evidence was found for orthographic markedness effects, albeit with a different pattern for word and non-word processing: while marked words were facilitated (responded to faster and more accurately) compared to unmarked words, the opposite pattern emerged for non-words. The pattern of results was comparable in both proficiency groups. No evidence emerged for the influence of first language (L1) neighborhood on L2 word or non-word processing. Thus, the results emphasize the need to integrate orthographic markedness as a relevant psycholinguistic variable in bilingual models of visual word recognition such as BIA/+ and to take it into account when investigating cross- language effects and the issue of language non-selectivity during visual word recognition.

Keywords Visual word recognition · Bilingual · Orthographic neighborhood

Introduction

In the bilingual visual word recognition field, the language non-selectivity lexical access hypothesis assumes that all lexical representations that share some form (e.g., orthographic) overlap with a visual input are automatically co-activated, whatever the language they belong to. This mechanism has been integrated into the Bilingual Interactive Activation theoretical framework (BIA/+, Dijkstra & van Heuven, 2002; van Heuven, Dijkstra, & Grainger, 1998) both derived from the McClelland and Rumelhart (1981) monolingual IA model. It predicts that lexical representations from both languages of a bilingual individual are connected through inhibitory connections; as a consequence, when a visual input is presented, both within- and cross-language interference can arise from lexical candidates that are orthographically close to the target input.

As pointed out by van Heuven et al. (1998), this hypothesis needs to be tested in a pure 'monolingual mode' that is in the absence of words from the non-target language. Hence, examination of orthographic neighborhood effects across languages are of particular interest. Orthographic substitution neighbors are all words that share all letters but one with a target word, at the same position; these neighbors are within or across languages (e.g., *fire* has neighbors such as *hire* in English and *dire*, say in French). Yet, despite strong theoretical significance of these effects for testing bilingual models' predictions, only a few studies have directly addressed this issue (Bijeljac-Babic, Biardeau, & Grainger, 1997; Dijkstra, Hilberinck-Schulpen, & van Heuven, 2010; Dirix, Cop, Drieghe, & Duyck, 2017; Midgley, Holcomb, van Heuven, & Grainger, 2008; van Heuven et al., 1998). Using the masked priming paradigm, Bijeljac-Babic et al. (1997) firstly uncovered a cross-language orthographic neighborhood frequency effect:

✉ Eva Commissaire
commissaire@unistra.fr

[1] Laboratoire de Psychologie des Cognitions (LPC-EA4440), Faculté de Psychologie, Université de Strasbourg, 12 rue Goethe, 67000 Strasbourg, France

[2] SCALab (UMR CNRS 9193), Université de Lille, rue du Barreau, 59653 Villeneuve d'Ascq, France

French/English bilinguals took longer to recognize target words when preceded by highly frequent orthographically related prime words from the non-target language (e.g., *game-GAGE*, forfeit) compared with unrelated words (e.g., *bird-GAGE*). This effect, which was found both from L1-to-L2 and L2-to-L1, was thought to reflect inhibitory links among lexical representations from the two languages (see also Dijkstra et al., 2010). Using lexical decision and perceptual identification tasks in a pure monolingual mode, van Heuven et al. (1998; see also ERP study by Midgley et al., 2008) showed that Dutch words with a large number of within- *and* cross-language (English) neighbors were recognized more slowly than words with a small neighborhood. This result was recently replicated by Dirix et al. (2017), albeit in a generalized lexical decision task that requires deciding the language membership of a word, and in natural sentence reading (yet leading to facilitatory effects). Again, these neighborhood effects were hypothesized to reflect language non-selective lexical access.

Although these effects are taken for granted, alternative interpretations remain to be discussed and tested. Our position is that the effect of orthographic neighborhood needs to be dissociated from the effect of a sub-lexical variable, recently referred to as "orthographic markedness" (Casaponsa, Carreiras, & Duñabeitia, 2014; Casaponsa & Duñabeitia, 2015; van Kesteren, Dijkstra, & de Smedt, 2012). Indeed, L2 words can contain orthographic patterns that are shared across languages (e.g., *house* for a French speaker), or else very specific to that language due to the presence of L2-specific letters, bi/trigrams or graphemes (e.g., *right*), referred to as marked words. If uncontrolled, it is highly likely that words with a large cross-language neighborhood mostly refer to words with shared sub-lexical orthographic patterns across languages, whereas words with a small neighborhood are rather composed of more L2-specific orthographic patterns. Moreover, orthographic markedness seems to impact bilingual word recognition per se. Language cues such as specific letters or bigrams would help to determine the language membership of a word (Casaponsa et al., 2014; van Kesteren et al., 2012). In addition, language non-selectivity during lexical access could be reduced for marked words that contain language-specific bigrams; that is, connections between L1 and L2 could be more likely observed for words *without* L2-specific patterns (Casaponsa & Duñabeitia, 2015). So far, whether lexical access differs for marked versus unmarked words, while controlling for cross-language orthographic neighborhood, remains unclear.

The goal of the present study was to jointly examine the role of these lexical and sub-lexical variables on L2 visual word recognition. Two groups of native French speakers varying on English (L2) proficiency performed an English lexical decision task with three word and non-word conditions: (1) L2-specific orthography combined with small cross-language orthographic neighborhood size (Marked small N condition), (2) language-shared orthography with small cross-language orthographic neighborhood size (Unmarked small N condition), and (3) with large cross-language orthographic neighborhood size (Unmarked large N condition). Based on van Heuven et al. (1998), longer reaction times were expected for the large N compared to the small N conditions (see also Lemhöfer & Radach, 2009, on non-word processing). Such an effect would reflect the inhibitory influence of French lexical representations, their dominant language, on the recognition of English words (and rejection of non-words), and thus language non-selectivity during lexical access. Whether the pattern of results would be affected by the distinction between marked versus unmarked (non) words was unclear. In addition, how proficiency might impact these patterns of results was explored.

Method

Participants

A total of 73 native French speakers participated in the study, among whom 41 were considered as low proficient in English (L2), and 32 as highly proficient. They were recruited at the University of Strasbourg, France. None of the participants had ever lived in an English-speaking country nor had been raised in a bilingual environment. They all had learned English at secondary school where exposure to English is around 3–4 h per week but differed according to their L2 proficiency, which was measured through subjective ratings and objective measures of L1-to-L2 translation (Casalis, Commissaire, & Duncan, 2015).

Stimuli

A total of 90 4- to 5-letter long monosyllabic English words (see Appendix 1) were selected and classified into three conditions: (1) Marked small N referred to English words that contained at least one very unlikely bigram according to French orthography (e.g., letter sequences such as wr- or -ck as in the word *wreck*) and, consequently, a small cross-language orthographic neighborhood; (2) unmarked small N corresponded to English words with few cross-language orthographic neighbors (and with low frequency[1]) but whose orthographic constituents were more legal in French compared to the specific condition (e.g., the word *spare*); and (3) unmarked large N referred to English words with legal orthographic patterns according to French but also a larger cross-

[1] Both neighborhood size and frequency (of the most frequent neighbor) were taken into account to select the stimuli given on-going debates on which of these variables is the most relevant psycholinguistic factor in monolinguals.

language orthographic neighborhood (e.g., *pride* has seven French neighbors such as *prise*, plug). Table 1 presents the linguistic features of these word conditions.

These conditions differed on cross-language *sub-lexical* and *lexical* orthographic features, which were estimated by using the Lexique database (New et al., 2001). We used minimal bigram frequency according to French orthography[2] to contrast marked from unmarked conditions (Westbury & Buchanan, 2002). The conditions differed with respect to French minimal bigram frequency, $F(2,87) = 55.579$, $p < .001$, $\eta^2_p = .56$; this reflected lower frequency in the marked compared to both unmarked conditions ($p < .001$ for both comparisons), which did not differ from each other ($p = .15$, n.s.). They also differed on neighborhood *size*, $F(2,87) = 130.88$, $p < .001$, $\eta^2_p = .75$, and *frequency* of the most frequent French neighbor, $F(2,87) = 22.293$, $p < .001$, $\eta^2_p = .34$. English words from the unmarked large N condition had more French neighbors compared to unmarked and marked small N conditions ($p < .001$ for both comparisons), which did not differ from each other ($p = .34$, n.s.). Their most frequent French neighbor was also on average more frequent in the unmarked large N compared to both small N conditions ($p < .001$ for both comparisons), which again did not differ from each other ($p = .99$, n.s.). The three conditions were matched on many other dimensions: number of letters, $F < 1$, n.s., and of phonemes, $F(2,87) = 1.345$, $p = .27$, n.s, lexical printed frequency (occurrences per million), $F < 1$, n.s, English bigram frequency (irrespective of word length and bigram position), $F(2,87) = 1.075$, $p = .30$, n.s., but also English orthographic neighborhood size, $F(2,87) = 1.59$, $p = .20$, n.s., and frequency (occ. per million), $F < 1$, n.s. (taken from the MCWord database, Medler & Binder, 2005[3]).

Pseudoword items (see Appendix 2) were constructed by changing one letter for target words (see Table 2). The conditions differed on French minimal bigram frequency, $F(2,87) = 22.602$, $p < .001$, $\eta^2_p = .34$, reflecting lower frequency in the marked compared to both unmarked conditions ($p < .001$ for both comparisons), but no difference between these two ($p = .89$, n.s.). They also differed on neighborhood *size*, $F(2,87) = 32.347$, $p < .001$, $\eta^2_p = .43$, and *frequency* of the most frequent French neighbor, $F(2,87) = 9.331$, $p < .001$, $\eta^2_p = .18$. Pseudowords from the unmarked large N condition had more French neighbors compared to unmarked and marked small N conditions ($p < .001$ for both comparisons), which were comparable ($p = .19$, n.s.). Their most frequent French neighbor was also on average more frequent in the unmarked large N condition compared to both small N conditions ($p < .01$ for both comparisons), which again did not differ from each other ($p = .99$, n.s.). The pseudoword conditions were based on word stimuli and were thus matched on letter and phoneme length, as well as on English bigram frequency, $F(2,87) = 1.842$, $p = .16$, n.s., orthographic neighborhood size, $F(2,87) = 1.773$, $p = .18$, n.s., and frequency, $F < 1$, n.s.

Procedure

A fixation point was presented for 500 ms and followed by the target item that remained on screen until the participant's response or else for 3,000 ms. Participants were asked to decide as quickly and accurately as possible whether the visual item was a real English word or not. "Yes" responses were performed using their dominant hand.

Results

Several words (*blur*, *chore*, *merge*, *rouse, knot*, and *swap*) and one pseudoword (*smale*) had poor responses (less than 70% accuracy) and were removed from further analyses. All data points below 300 ms and above 2,000 ms were removed from reaction time analyses, i.e., less than 1% and 2.5% of word and pseudoword data, respectively. Participants' means (and standard deviations) by group and for all conditions are shown in Table 3. Reaction times (inverse transformed) and accuracy data on both words and pseudowords were respectively examined by using linear mixed models and binomial mixed models (by using the lme4 package on R, Bates, Mächler, Bolker, & Walker, 2015). For all the models considered below, we included a random intercept that allows considering the structure of our data (dependencies among measures obtained by the same individual) and the inter-individual performance variability.[4] Note that items were not entered as a random factor given the strong constraints on stimuli selection (exhaustivity); however, we also provide information regarding the inclusion of such a variable as it informs on by-item variability issues (see Discussion). We also used a model comparison approach by examining (1) the Akaike Information Criteria (AIC) according to which lower values equal a better model fit to data (Akaike, 1973) and (2) *p*-values based on likelihood ratio test comparisons.

Firstly, we examined the influence of the cross-language orthographic neighborhood by comparing a model (model 1) that assumed a performance difference between the three conditions with a model (model 2) that constrained the mean performances in the "small N" and "large N" conditions to be equal (the markedness effect being kept constant

[2] Minimal bigram frequency refers to the frequency of the least frequent bigram of a word, here from a cross-language approach (e.g., the bigram wr- from the word *wreck*, which does not occur in French orthography).

[3] Items were also matched on lexical frequency based on the Subtlex-UK database (van Heuven et al., 2014), which provides subtitles' frequencies.

[4] We also added a random slope to the model that examined whether the condition effect may vary between participants. It never provided better adjustment and it is thus not reported.

Table 1 Linguistic features of English (L2) word stimuli according to condition

	Marked small N	Unmarked small N	Unmarked large N	Effect
Number of letters	4.43 (.5)	4.5 (.51)	4.43 (.5)	n.s.
Number of phonemes	3.27 (.58)	3.5 (.57)	3.33 (.55)	n.s.
Lexical frequency	149 (271)	162 (325)	161 (343)	n.s.
English N-size	7.2 (4.66)	7.1 (4.36)	8.97 (4.53)	n.s.
English N-frequency	41 (64)	114 (402)	50 (59)	n.s.
English bigram frequency	2775 (1308)	3172 (1362)	3203 (1121)	n.s.
French (L1) N-size	.93 (.94)	1.43 (1.14)	6.1 (1.85)	< .001
French (L1) N-frequency	8.02 (10.34)	4.88 (5)	212.3 (244.3)	< .001
French (L1) minimal bigram frequency	73 (114)	2251 (1166)	2761 (1388)	< .001

elsewhere). Similarly, we examined the influence of markedness by comparing model 1 with a model (model 3) that constrained the mean performances in the "marked" and "unmarked" conditions to be equal (the neighborhood effect being kept constant elsewhere).

Secondly, we compared the best previous fitted model to a model that (1) also included the fixed effect of group (model 4) and (2) that added the fixed effect of group and the fixed effect of interaction between group and the other variable (s) having a significant fixed effect (model 5).

Word data

Reaction times

While model 1 (AIC = -74,079) had a better fit than model 3 (AIC = -74,070, $X^2(1, 73) = 10.56$, $p<.01$), suggesting an effect of orthographic markedness, no difference emerged between models 1 and 2 (AIC = -74,080, $X^2 < 1$, n.s.). Reaction times were indeed faster for the small N marked condition (678 ms) compared to both unmarked conditions, which were numerically very close to each other (696 ms and 693 ms for small N and and large N unmarked conditions, respectively).

When adding to the markedness model the fixed effect of group, the resulting model 4 was better (AIC = -74,087; $X^2(1, 73) = 8.68$, $p<.01$) but model 5, which included an interaction effect between markedness and group, did not improve the model adjustment (AIC = -74,085, $X^2 < 1$, n.s.). Indeed, the highly proficient group responded much faster (651 ms) compared to the low proficiency group (721 ms), but the overall pattern was sensibly comparable among the two groups (see Table 3).

Errors

Model 1 (AIC = 3,096) had a better fit than model 3 (AIC = 3,109, $X^2(1, 73) = 15.09$, $p<.001$), suggesting an effect of orthographic markedness; again, no difference emerged between models 1 and 2 (AIC = 3,094, $X^2 < 1$, n.s.). Participants made less errors on the small N marked condition (6.1% errors) compared to both small N and large N unmarked conditions (9.1% errors on both conditions).

Adding the fixed effect of group to the orthographic markedness model (model 4) improved the model adjustment4 (AIC = 3,086; $X^2(1, 73) = 9.69$, $p<.01$), but model 5 (AIC = 3,088) did not improve the adjustment. Accuracy was much higher in the highly proficient group (4.4% errors) than in the lower proficiency group (11.1%).

Pseudoword data

Reaction times

Model 1 (AIC = -77123) had a better fit than model 3 (AIC = -77,108, $X^2(1, 73) = 16.69$, $p<.001$), suggesting an effect of orthographic markedness; again, no difference emerged between models 1 and 2 (AIC = -77,125, $X^2 < 1$, n.s.).

Table 2 Linguistic features of English (L2) pseudoword stimuli according to condition

	Marked small N	Unmarked small N	Unmarked large N	Effect
English N-size	4.67 (4.71)	5.63 (3.78)	7.73 (4.42)	n.s.
English N-frequency	87.84 (156)	58.82 (109.3)	110.14 (256)	n.s.
English bigram frequency	2454 (1068)	3069 (1392)	2935 (1429)	n.s.
French (L1) N-size	.63 (1.19)	1.47 (1.91)	4.27 (2.24)	< .001
French (L1) N-frequency	6.41 (17.21)	12.77 (22.49)	206.45 (352)	< .001
French (L1) minimal bigram frequency	52 (62)	2308 (2167)	2141 (1262)	< .001

Table 3 Reaction times (in ms) and percentages of errors for English (L2) word and pseudoword items, according to group and condition

	Marked small L1 N	Unmarked small L1 N	Unmarked large L1 N
Word items			
Highly proficient group			
Reaction times	644 (122)	656 (120)	660 (24)
Errors	3.4 (4)	4.8 (5)	4.9 (7)
Low proficiency group			
Reaction times	714 (135)	739 (136)	734 (146)
Errors	8.3 (10)	12.5 (12)	12.4 (12)
Pseudoword items			
Highly proficient group			
Reaction times	885 (204)	857 (211)	851 (196)
Errors	15.1 (13)	8.9 (10)	9.5 (10)
Low proficiency group			
Reaction times	872 (202)	836 (167)	834 (171)
Errors	12.8 (11)	7.8 (8)	7.9 (9)

Pseudoword rejection took longer for the small N marked condition (862 ms) compared to both unmarked conditions, small N (838 ms) and large N (832ms).

No better model fit was found when adding to the markedness model the fixed effect of group (model 4, AIC = -77,123) or the interaction between markedness and group (model 5, AIC = -77,121). Indeed, comparable rejection times were observed in the two participant groups (854 ms and 836ms for the high and low proficiency groups, respectively).

Errors

Model 1 (AIC = 3913) had a better fit than model 3 (AIC = 3,948, $X^2(1, 73) = 36.32$, $p<.001$), suggesting an effect of orthographic markedness; no difference emerged between models 1 and 2 (AIC = 3,911, $X^2 < 1$, n.s.). Participants made more false alarms on the small N marked condition (13.8% errors) compared to both small N and large N unmarked conditions (8.3% and 8.6% errors, respectively).

The orthographic markedness model remained the best fitted model when compared with model 4, which included the fixed effect of group (AIC = 3913), and model 5, which included both main fixed effect of group and interaction between the two fixed effects (AIC = 3,915). Indeed, comparable false alarms were observed in the two groups (11.2% and 9.5% errors for the high and low proficiency groups, respectively).

Discussion

This study investigated the extent to which lexical access in L2 is influenced by both L1 orthographic neighborhood and markedness. Only orthographic markedness had an impact on L2 processing and no effect of cross-language orthographic neighborhood emerged in the present study. Interestingly, orthographic markedness in L2 led to a facilitation effect for word items (on reaction time and accuracy) but an inhibition one for non-word items, independent of participants' L2 proficiency. Importantly, statistical analyses on reaction times available from the British Lexicon Project (Keuleers, Lacey, Rastle, & Brysbaert, 2012) showed no effect of either orthographic variable in native English speakers, confirming the impact of French (L1) knowledge. Van Kesteren et al. (2012) showed that marked bigrams could directly provide information about language membership of items (language nodes); however, this information is used by the task/decision system (BIA+, Dijkstra & van Heuven, 2002) only when it provides diagnostic information to correctly perform a specific task. In the present study, though, both words *and* nonwords were either L2 marked or unmarked and the task was in L2 only, excluding any L1 word. As markedness effects still emerged, this suggests that it has an impact on the word recognition system per se. We hypothesize that marked words would likely reduce the activation of the non-target language, leading to faster recognition (see Casaponsa & Duñabeitia, 2015), while marked non-words would add difficulty to the rejection decision due to stronger activation of the target language membership. Note that when we included by-item variability in the statistical models, the markedness effect decreased in the word data analyses but not in the pseudoword data analyses. Other sources of item variability may have affected L2 word processing in the current study despite great controls in stimuli selection and follow-up checks of adequate matching on item' familiarity for English as L2 speakers (based on subjective ratings provided by an independent sample of late learners of English).

Surprisingly, we found no evidence for cross-language interference at the lexical level. While other psycholinguistic factors such as language or lexical frequency could interact with neighborhood effects, it is noteworthy that the reaction times were numerically very close between small N and large N (when matched on markedness) in both proficiency groups and for both words and non-words. Our data are in line with the conclusions by Lemhöfer et al. (2008), who report no impact of cross-language orthographic neighborhood on L2 word recognition time. However, they differ from those of van Heuven et al. (1998) and Dirix et al. (2017). To us, this discrepancy demonstrates that cross-language neighborhood effects could have possibly been confounded with markedness effects.

To conclude, our study highlights the need to specify cross-language influences at the lexical and sub-lexical levels. Doing so, we found evidence of an effect of orthographic markedness but not of cross-language neighborhood, an index of language non-selectivity at the lexical level. Future studies will help to clarify the mechanisms underlying orthographic markedness effects in order to further integrate these sub-lexical orthographic mechanisms (see also Commissaire, Duncan, & Casalis, 2014, on L2 grapheme processing) into the BIA+ model.

Appendix 1

Table 4 List of word items

Marked small N	Unmarked small N	Unmarked large N
BRAKE	BLADE	CRAVE
SHARE	CHEST	FRAME
SHINE	CURVE	HORSE
STRAW	PROUD	MERGE
SWEAR	NURSE	PASTE
THREE	SCALE	PRICE
THROW	SCARE	ROUSE
TWICE	SCOPE	SINCE
WATCH	SPARE	TASTE
WITCH	START	TRADE
WORSE	STONE	CHORE
WRECK	CURSE	PRIDE
WRITE	DRIVE	TENSE
COOK	MOUSE	BEND
HIGH	HOUSE	CUTE
KNOT	BEST	DARE
LEAK	BIND	DIVE
LINK	BLUR	FAIL
MONK	FIST	FATE

Table 4 (continued)

Marked small N	Unmarked small N	Unmarked large N
PINK	GLAD	LOSE
TOWN	HAND	NOSE
STAY	HOME	RULE
SWAP	HOPE	SAIL
TALK	JAIL	SAME
TWIN	LAND	SOIL
BACK	HOLE	TIME
BLOW	NAIL	FIVE
GREY	THAN	JUNE
WRAP	CLUE	SIDE
TAKE	SCAR	NAME

Appendix 2

Table 5 List of pseudoword items

Marked small N	Unmarked small N	Unmarked large N
FRAKE	BURVE	CHORL
SBRAW	CHIST	CRIDE
SHERE	CLADE	FRADE
SHREE	CRIVE	FRAVE
SHROW	HOUBE	FRUME
SHUNE	MOUFE	HORTE
SWEAL	NARSE	JASTE
TWIRE	SCOME	LERGE
WALCH	SLARE	MASTE
WILCH	SLART	NOUSE
WORSH	SMALE	PRINE
WRILE	SPURE	TINCE
WRICK	STINE	TUNSE
DACK	TROUD	BAME
PAKE	TURSE	DARM
DLOW	BLOR	FAIP
HAGH	BIST	FIME
KNAT	BINF	FOSE
NINK	CLIE	FUNE
MALK	DIST	LOIL
MEAK	GLAT	MIVE
PENK	HANB	MOSE
PONK	HOLT	NAFE
POOK	LAIL	NEND
STAW	LANF	SAIB
SWAL	NAUL	SIVE
TAWN	SMAR	SODE
TWAN	TRAN	SULE
VREY	HOFE	VATE
WRIP	VOPE	VUTE

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Caski (Eds.), Proceedings of the Second International Symposium on Information Theory (pp. 267–281). Budapest: Akademiai Kiado.

Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1-48.

Bijeljac-Babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory and Cognition, 25, 447-457.*

Casalis, S., Commissaire, E., & Duncan, L. (2015). Sensitivity to morpheme units in English as L2 word recognition. *Writing System Research, 7*(2), 186-201.

Casaponsa, A., Carreiras, M., & Duñabeitia, J. A. (2014). Discriminating languages in bilingual contexts: the impact of orthographic markedness. *Frontiers in Psychology*, *5*, 424.

Casaponsa, A., & Duñabeitia, J. A. (2015). Lexical organization of language-ambiguous and language-specific words in bilinguals. *The Quarterly Journal of Experimental Psychology*, *69*(3), 589-604.

Commissaire, E., Duncan, L. G., & Casalis, S. (2014). Grapheme coding in L2: how do L2 learners process new graphemes? *Journal of Cognitive Psychology, 26*(7), 725-739.

Dijkstra, A., Hilberink-Schulpen, B., & van Heuven, W.J.B. (2010). Repetition and masked form priming within and between languages using word and non-word neighbors. *Bilingualism: Language and Cognition, 13*, 341-357.

Dijkstra, A., & 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.

Dirix, N., Cop, U., Drieghe, D., & Duyck, W. (2017). Cross-lingual neighborhood effects in generalized lexical decision and natural reading. *Journal of Experimental Psychology, Learning, memory, and cognition*, *43*(6), 887-915.

Keuleers, E., Lacey, P., Rastle, K., & Brysbaert, M. (2012). The British Lexicon Project: lexical decision data for 28,730 monosyllabic and disyllabic English words. *Behavioral Research*, *44*, 287-304.

Lemhöfer, K., Dijkstra, A., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008) Native language influences on word recognition in a second language: a megastudy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 12-31.

Lemhöfer, K., & Radach, R. (2009). Task context effects in bilingual nonword processing. *Experimental Psychology*, *56*(1), 41-47.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological review*, *88*(5), 375-407.

Medler, D.A., & Binder, J.R. (2005). MCWord: An On-Line Orthographic Database of the English Language. http://www.neuro.mcw.edu/mcword/

Midgley, K. J., Holcomb, P. J., Van Heuven W. J. B., & Grainger, J. (2008). An electrophysiological investigation of cross-language effects of orthographic neighborhood. *Brain Research, 1246*, 123-135.

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.

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

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

Van Kesteren, R., Dijkstra, T., & de Smedt, K. (2012) Markedness effects in Norwegian–English bilinguals: Task-dependent use of language-specific letters and bigrams. *The Quarterly Journal of Experimental Psychology*, *65*, 2129-2154.

Westbury, C., & Buchanan, L. (2002). The probability of the least likely non-length controlled bigram affects lexical decision reaction times. *Brain and Language*, *81*(1-3), 66-78.