How do bilinguals identify the language of the words they read?

Article *in* Brain research · July 2015

DOI: 10.1016/j.brainres.2015.07.035 · Source: PubMed

3 authors:

Aina Casaponsa
Lancaster University

Manuel Carreiras
Basque Center on Cognition, Brain and Language

Jon Andoni Duñabeitia
Basque Center on Cognition, Brain and Language

How do bilinguals identify the language of the words they read?

Aina Casaponsa[1], Manuel Carreiras[1,2] and Jon Andoni Duñabeitia[1]

[1] BCBL. Basque Center on Cognition, Brain and Language; Donostia, Spain

[2] Ikerbasque, Basque Foundation for Science; Bilbao, Spain

Contact information:

Aina Casaponsa Galí

Basque Center on Cognition, Brain and Language (BCBL)

Paseo Mikeletegi 69, 2nd floor

20009 Donostia, SPAIN

a.casaponsa@bcbl.eu

Telephone: +34943309300

Abstract

How do bilinguals detect the language of the words they read? Recent electrophysiological research using the masked priming paradigm combining primes and targets from different languages has shown that bilingual readers identify the language of the words within approximately 200 milliseconds. Recent evidence shows that language-detection mechanisms vary as a function of the orthographic markedness of the words (i.e., whether or not a given word contains graphemic combinations that are not legal in the other language). The present study examined how the sub-lexical orthographic regularities of words are used as predictive cues. Spanish-Basque bilinguals and Spanish monolinguals (control group) were tested in an Event-Related Potential (ERP) experiment, using the masked priming paradigm. During the experiment, Spanish targets were briefly preceded by unrelated Spanish or Basque words. Unrelated Basque words could contain bigram combinations that are either plausible or implausible in the target language (Spanish). Results show a language switch effect in the N250 and N400 components for marked Basque primes in both groups, whereas, in the case of unmarked Basque primes, language switch effects were found in bilinguals but not monolinguals. These data demonstrate that statistical orthographic regularities of words play an important role in bilingual language detection, and provide new evidence supporting the assumptions of the BIA+ extended model.

Keywords: Bigrams, Bilingualism, Multilingual reading, Orthographic cues, Masked language switch cost priming, Event-related potentials.

1. Introduction

It is well-established that when bilinguals process words, they unconsciously access both of their languages, even when set in a single language context (Dijkstra et al., 2000; Dimitropoulou et al., 2011a, 2011b; Duyck et al., 2007; Duyck and Warlop, 2009; Grossi et al., 2012; Hoshino and Thierry, 2012; Ng and Wicha, 2013; Midgley et al., 2008; Perea et al., 2008; Spalek et al., 2014; Schwartz et al., 2007; Thierry an Wu, 2004; Thierry and Wu, 2007; Van Heuven et al., 1998, 2008; Zhang et al., 2011). However, bilinguals need to know in which language they are reading to correctly retrieve the meaning of words. Consider the case of false friends (words that strongly overlap between languages but have different meanings). For instance, the word *pie* refers to a culinary preparation encased in pastry in English and ′foot' in Spanish, which cannot be disambiguated without knowing in which language the word is presented (i.e., language attribution).

Despite the importance of language attribution in bilinguals, the mechanisms by which a given word form is associated with a given language remain unclear. The main aim of this study is to provide new evidence about the mechanisms underlying language identification by exploring the time-course of unconscious and automatic language switch effects in reading. More specifically, we investigated whether language-specific sub-lexical cues can drive language attribution.

Several studies have suggested that in an ambiguous language context (i.e., when languages have similar scripts; e.g., Spanish and Catalan) language information is accessed through the lexical representations of words (Chauncey et al, 2008, 2011; Dijkstra et al., 1999; Dijkstra and Van Heuven, 2002; Midgley et al., 2009a, 3009b; Von Studnitz and Green, 2002). In this case, effects of language identification result from top-down modulations from the language nodes feeding information back to the lexical units. However none of these studies have explored whether these effects can be modulated by sub-lexical information.

In many language combinations sharing the same script (e.g., Dutch and English or Basque and Spanish), one can easily find words that are orthographically marked (i.e., words that contain sub-lexical features that are only plausible in one language). For example, the Basque word *neska* -´girl´, contains the bigram "sk" which is illegal in Spanish. At the same time, it is relatively easy to find words that are orthographically unmarked, such as the Basque word *mutil* -´boy´, which complies with Spanish orthotactics[1]). Along these lines, Van Kesteren et al. (2012) showed that language-specific orthography can guide language decisions in bilinguals, and proposed a direct link between sub-lexical information and language membership. In their study, Norwegian-English bilinguals completed a series of language decision tasks featuring marked and unmarked Norwegian and English words. The authors demonstrated that language membership can be accessed via lexical representations, but critically also via sub-lexical levels (see also Casaponsa et al., 2014; Casaponsa and Duñabeitia, in press; Lemhöfer et al., 2011; Orfanidou and Sumner, 2005; Vaid and Frenck-Mestre, 2002). Therefore, Van Kesteren et al. (2012) proposed an extension to the Bilingual Interactive Activation Plus model (BIA+ ; Dijkstra and Van Heuven, 2002). The addition to the model is a sub-lexical language node, which can be accessed directly through excitatory connections from sub-lexical levels and then be read out by the task/decision system. Thus, an orthographically marked word containing language-specific orthographic cues can lead to effective language attribution solely on the basis of sub-lexical information. Although the number and nature of the connections of the sub-lexical language node with other levels of word processing remain underspecified, the BIA+ extended model highlights the importance of language-specific orthography in reading (see Casaponsa and Duñabeitia, in press, for review).

[1] Orthotactics is a term used to define the orthographic characteristics of the existent vocabulary in a specific language that focuses on the sequences of letter combinations that are allowed in that language.

The present study aims to shed light on the influence of orthographic markedness by tracking the time-course of language identification in bilingual word recognition, by investigating the types of cues that help bilingual readers detect the language code. The masked priming language-switching paradigm combined with Event-Related Potential (ERP) recording is an extremely valuable approach for researchers exploring the time-course of language identification. Using such approach, it has been shown that bilinguals can automatically and unconsciously detect the language of a word within 200 milliseconds of stimulus onset (Chauncey et al., 2008, 2011; Dunabeitia et al., 2010b). In masked priming ERP studies exploring language detection, participants are typically presented with primes for 50 ms (i.e., below the threshold of conscious recognition) and unrelated targets that are in the same or different languages (e.g., casa-DOG vs. house-DOG, where *casa* means ‘house’ in Spanish). This way, the only relationship between targets and primes is the language code, which allows studying the cost associated with an unconscious language switch.

One of the first studies exploring this language switch effect using an electrophysiological masked priming paradigm was that of Chauncey et al. (2008) testing French-English bilinguals. They found that switch trials elicited more negative-going waveforms than non-switch trials in the windows of two components, the N250 and N400. The authors suggested that these increased negativities for switch compared to non-switch trials resulted from automatic top-down modulation from the language node feeding back to lexical representations. Thus, the language nodes of the prime words are presumed to inhibit the target language lexical representations in switch trials and/or to enhance the activation of the same-language lexical representations in non-switch trials (see also Duñabeitia et al., 2010b). Other ERP studies have shown similar patterns (Martin et al., 2012; Midgley et al., 2009a; Proverbio et al., 2004; Van Der Meij et al., 2011), but to date no ERP study has investigated the contribution of orthographic cues.

However, the time-course of visual word recognition based on results from experiments using masked priming relies on the interpretation of the N250 as typically associated with the mapping of orthographic units onto orthographic word forms (see Grainger & Holcomb, 2009, for a review), and the N400 with a later stage of processing sensitive to interactions between lexical forms and semantic representations. Therefore, the former component reflects sub-lexical and lexical information integration while the later reflects lexical-semantic integration. Thus it is possible that bilinguals use both lexical and sub-lexical routes to access language membership, giving rise to the early effects found in masked language-switching priming paradigm (in line with the parallel language identification routes sketched by Van Kesteren et al., 2012). Along these lines, one could predict that under masked priming conditions in a seemingly monolingual context, a fast-operating sub-lexical route highly sensitive to cross-language orthographic regularities would initially detect conflicting graphemic chunks in the masked primes from a different language. This is consistent with the study by Casaponsa and Duñabeitia (in press) in which marked masked primes did not yield significant lexico-semantic effects on target word processing. Despite the theoretical and experimental plausibility of this hypothesis, to our knowledge no masked language-switching priming experiment has manipulated orthographic markedness. In fact, masked language-switching priming experiments published so far have used cross-script priming, which warrants the presence of sub-lexical cues (e.g., Hoshino et al., 2010; Dimitropoulou et al., 2011b), or same-script priming without controlling for orthographic markedness (e.g., Chauncey et al., 2008; Dimitropoulou et al., 2011a, 2011b; Midgley et al., 2009b, Perea et al., 2008). Here we explore the effects associated with sub-lexically marked and unmarked switch trials and compared them to non-switch trials.

In a nutshell, the current study explores how highly proficient Spanish-Basque bilinguals access language information about the words they read. Our guiding hypothesis is that bilinguals do not rely solely on lexical-level information but also on sub-lexical statistical regularities. We used the ERP masked language-switching priming paradigm while participants performed a go/no-go semantic categorization task. Target words were always in the dominant language (L1-Spanish), and the masked primes were either in the dominant or in the non-dominant language (L1 or L2). Critically, language-specific orthographic markedness of the bigrams of L2 primes was manipulated as a function of the L1 (target) orthography to explore how and when sub-lexical information modulates language identification.

According to the dual-route model of language identification (BIA+ extended, Van Kesteren et al., 2012), if within-word statistical regularities guide bilingual language detection, then Basque (L2) primes that are formed by Basque-specific marked bigrams should modulate the masked language switch cost effect in initial stages of target word processing (i.e., N250) and produce greater switch costs as compared to masked primes that follow the orthotactic rules of the target language (i.e., unmarked Basque primes). Furthermore, under the assumption that sub-lexical information is not informative enough to detect unmarked primes as belonging to Basque, language switch effects are only expected (if any) at later stages of processing (e.g., in the N400 window; see Chauncey et al., 2008).

In addition to the bilingual group tested in Experiment 1, a second experiment was conducted with a control group of Spanish monolinguals without any prior knowledge of Basque using the same materials and procedure (Experiment 2). This monolingual control group was tested in order to ensure that the differential masked language switch effects found for orthographically marked and unmarked words in the bilingual group could not be accounted for by the saliency of the sub-lexical orthotactics of the marked items. We expected that marked Basque prime words would be processed as non-words (i.e., strings containing illegal combinations that cannot be recognized as potential words), while unmarked Basque prime words were expected to act as pseudo-words (i.e., strings that lack a referent in the mental lexicon even though they follow their language orthotactics). Therefore, marked Basque primes were expected to elicit greater negativities than unmarked Basque primes, based on sub-lexical feature analysis (see Grossi & Coch, 2005). At the same time, we expected Basque marked and unmarked primes to elicit similar effects in the N400 window due to the absence of a lexical referent in Spanish (e.g., Grainger & Holcomb, 2009; Holcomb & Grainger, 2006). For the same reason, Basque words should elicit greater negativities than Spanish control words at this later stage of processing.

2. Results

2.1 Experiment 1: Spanish (L1) - Basque (L2) bilinguals.

2.1.1. Behavioral results

Bilingual participants correctly categorized 92.18% (SD=4.01) of the Spanish animal names when these words were presented as targets. Furthermore, when the animal names were presented as masked primes in the prime visibility test, none of the participants reported consciously perceiving them (or any other word), confirming that participants were unaware of the existence and nature of the masked primes (percentage of false alarms in the task: .59%, SD=.45; percentage detected primes in the prime visibility test: .47%, SD=2.18).

2.1.2. ERP results

The different language priming conditions split by bigram markedness are plotted in Figure 1 together with the voltage maps for the marked and unmarked switch effects in the different time windows of interest (see also Table 1 for means and standard deviations).

-Insert Table 1 around here-

2.1.2.1. 180-260 ms (N250)

The main effect of Language (switch vs. non-switch trials) was significant [$F(1,20)=7.48$, $p=.013$, $\eta^2_p=.27$], showing more negative-going waveforms for switch (i.e., L2→L1) than for non-switch trials (i.e., L1→L1). Critically, the interaction between Language and Bigram was significant [$F(1,20)=4.50$, $p=.047$, $\eta^2_p=.18$]. Planned comparison showed a modulation of the language switch cost effect as a function of the orthographic markedness of the Basque masked primes. A significant switch cost effect was found for marked primes compared to their L1 controls [$F(1,20)=11.67$, $p=.003$, $\eta^2_p=.37$], and not for the unmarked Basque primes [$F<1$, $p>.5$, $\eta^2_p=.02$] (see Figure 1). All other main effects and interactions did not approach significance [ps>.73].

2.1.2.2. 350-500 ms (N400)

The factors Language and Bigram significantly interacted with each other [$F(1,20)=5.80$; $p=.026$, $\eta^2_p=.23$]. Follow-up comparisons revealed a similar pattern to that found in the previous time window: a significant switch cost effect was found for the marked Basque primes [$F(1,20)=7.36$, $p=.013$, $\eta^2_p=.27$], but not for the unmarked Basque primes [$F<1$, $p>.8$, $\eta^2_p<.01$]. All other main effects and interactions did not reach significance [ps>.13].

-Insert Figure 1 around here-

2.1.3. Discussion

Bilingual participants showed a masked language switch cost effect in two main negative-going components (the N250 and N400) replicating and extending the findings reported in previous studies (Chauncey et al., 2008, 2011; Duñabeitia, Dimitropoulou, Uribe-Etxebarria, Laka, & Carreiras, 2010). We found a large and long-lasting switch cost effect only for those Basque words (L2) that were made of bigrams that were implausible in the L1 (Spanish). This masked language switch cost effect started as early as 200 milliseconds after the target word presentation (see Figure 1), and lasted for around 400 milliseconds. Critically, we did not find any significant switch cost effect for unmarked L2 words in any epoch (see Figure 1; see also Appendix for the visualization of the effects across the entire epoch).

2.2. Experiment 2: Spanish monolinguals

2.2.1. Behavioral results

Monolinguals correctly detected 91.84% (SD=4.25) of the animals (percentage of false alarms in the task: .38%, SD=.60). None of the participants reported consciously perceiving the animal names (or any other word) when these were presented as masked primes, confirming that participants were unaware of the existence and nature of the masked primes (percentage of primes detected in the prime visibility test= .32%, SD=1.00).

2.2.2. ERP results

The different language priming conditions split by bigram markedness are plotted in Figure 2 together with the voltage maps for the marked and unmarked effects in the different time windows of interest (see also Table 2 for means and standard deviations).

-Insert Table 2 around here-

2.2.2.1. 180-260 ms (N250)

The main effect of Language (switch vs. non-switch trials) was significant [F(1,20)=66.81, p<.001, η^2_p=.77], showing more negative-going waveforms for target words preceded by primes in the other language than for non-switch trials. This factor was partially modulated by the factor Region [F(2,40)=3.38, p=.07, η^2_p=.14, ε=.60]. Even though all three regions showed more negative-going waveforms for switch than non-switch trials, the effect sizes were greater over central and posterior regions [anterior: F(1,20)=25.69, p<.001, η^2_p=.56; central: F(1,20)=74.00, p<.001, η^2_p=.79; posterior: F(1,20)=80.57, p<.001, η^2_p=.80]. A marginal effect of Bigram was also found [F(1,20)=3.68, p=.07, η^2_p=.15], and the interaction between this factor and Language was marginally significant [F(1,20)=3.97, p=.06, η^2_p=.17], suggesting that the magnitude of the language effect was slightly modulated by the violation of L1 orthotactic rules (see Figure 2). Marked Basque masked primes elicited partially larger effects than unmarked Basque masked primes, but importantly these two conditions showed significant switch cost effects [$F(1,20)=38.51$, $p<.001$, $\eta^2_p=.66$, and $F(1,20)=11.34$, $p=.003$, $\eta^2_p=.36$, respectively]. All other interactions did not reach significance [ps>.35].

2.2.2.2. 350-500 ms (N400)

The main effect of Language was significant [$F(1,20)=13.01$, $p=.002$, $\eta^2_p=.39$], showing a general language switch cost effect. Switch trials were associated with more negative-going waveforms than non-switch trials. Critically, this effect was not modulated by the Bigram factor [interaction: $F<.1$, $p>.85$, $\eta^2_p<.001$], showing that in the N400 time window monolingual participants showed similar switch cost effects for marked and unmarked Basque primes. Besides, a marginal three way interaction was found [$F(2,40)=3.28$, $p=.07$, $\eta^2_p=.14$, $\varepsilon=.68$]. However, follow-up comparisons did not show differential Language effects as a function of Bigrams in any of the Regions [anterior: $F(1,20)=.84$, $p=.37$, $\eta^2_p=.04$; central: $F(1,20)=.24$, $p=.63$, $\eta^2_p=.01$; posterior: $F(1,20)=1.12$, $p=.30$, $\eta^2_p=.05$]. All other main effects and interactions did not result significant [ps>.31].

-Insert Figure 2 around here-

2.2.3. Discussion

Targets preceded by prime words in an unknown language showed more negative-going waveforms than targets preceded by words in the same language in both time windows (N250 and N400). Interestingly, this effect was slightly greater for non-word primes (marked Basque words) than for pseudo-word primes (unmarked Basque words) in the N250 time window. While the interaction in the N250 window only approached significance, it should be kept in mind that in sharp contrast with the results from the bilingual sample tested in Experiment 1, both conditions (orthographically marked and unmarked Basque primes) elicited larger negativities as compared to same-language conditions in Experiment 2. Besides, and again in contrast to the results from Experiment 1, robust N400 effects were found both for orthographically marked and unmarked Basque primes as compared to the Spanish control prime words. These results suggest that monolingual participants are also sensitive to L1 orthotactic violations at initial stages of visual word recognition. Moreover, these results show that monolingual readers are able to detect that a given word does not belong to their native language as soon as 200 ms, even when these strings follow the orthotactic rules of their language, demonstrating that some form of lexical access occurs within this epoch (see also Carreiras et al., 2009a, 2009b; Coch and Mitra, 2010; Duñabeitia et al., 2009; Grossi and Coch, 2005; Hoshino et al., 2010; Massol et al., 2011; Midgley et al., 2009b; Morris et al., 2007; Proverbio et al., 2009).

3. General discussion

The present study was designed to explore how bilinguals identify the language of the words they read. We investigated whether the magnitude of masked language switch cost effects is determined by language-dependent orthographic regularities. We aimed at exploring the influence of sub-lexical orthographic cues on the ERP patterns related to automatic and unconscious masked language switch effects that have been replicated recently with different bilingual samples in different languages (e.g., Chauncey et al., 2008; Dunabeitia et al., 2010a). To this end, highly proficient Spanish (L1) – Basque (L2) bilinguals were presented with L1 target words preceded by either unrelated L1 or L2 masked words (Experiment 1). Critically, the orthotactic plausibility in the L1 (Spanish; target language) of L2 masked prime words was manipulated. In addition to the bilingual group, a group of Spanish monolinguals was also tested for control purposes (Experiment 2).

The results showed a masked language switch cost effect in initial stages of reading (around 200 ms), which was also present at the later stages of visual word processing for Spanish-Basque highly proficient bilinguals. Importantly, these effects were modulated as a function of the orthographic markedness or specificity of Basque (L2) words for bilingual participants. They showed masked language switch effects only for those Basque words that were orthographically implausible in Spanish (i.e., orthographically marked items), and no effects for those Basque primes whose bigrams were plausible in Spanish orthography (i.e., orthographically unmarked items; see Appendix to better visualize these effects). Therefore, we demonstrated that bilingual readers unconsciously rely on basic orthographic cues in order to discriminate the words that do not correspond to the target language. Furthermore, results from Experiment 2 confirmed that this differential switch cost effect based on orthographic markedness is specific to bilinguals, since monolinguals showed strong effects for both unmarked and marked masked strings. These effects were slightly greater for strings that violated Spanish orthotactic regularities only at early stages of visual word processing (see Appendix).

Previous behavioral studies have shown that bilingual readers are highly sensitive to the bigram frequencies of words (Casaponsa, et al 2014; Casaponsa and Duñabeitia, in press; Grainger and Beauvillain, 1987; Thomas and Allport, 2000; Vaid and Frenck-Mestre, 2002; Lemhöfer et al., 2004, 2008, 2011; Van Kesteren et al., 2012; see also Grainger and Van Heuven, 2003; Dandurand et al., 2011; Hauk et al., 2006, 2008; Whitney and Cornelissen, 2008; Whitney et al., 2011, for evidence in monolinguals), and that orthographic information is explicitly used to identify the language of the words. Furthermore, recent data indicates that bilinguals use this information to speed up word recognition in ambiguous language contexts (e.g., Casaponsa et al., 2014) and to access language-selective or language-unspecific lexical representations (e.g., Casaponsa and Duñabeitia, in press), corroborating the hypothesis of two parallel routes to language identification (i.e., the lexical and sub-lexical routes; see Van Kesteren et al., 2012). The current study demonstrates that the manner in which the bilingual visual word recognition system detects the language of words is modulated by orthographic statistical information (i.e., orthotactic regularities) associated with the probabilistic combination of letters that are allowed in the context language(s). Furthermore, the current data suggest that language detection occurs at early automatic stages of visual word processing, even under circumstances in which the language code switch is not consciously perceived (i.e., under masked priming conditions), extending preceding evidence from conscious reading paradigms (Casaponsa et al., 2014; Grainger and Beauvillain, 1987; Lemhöfer et al., 2011; Vaid and Frenck-Mestre, 2002; Van Kesteren et al., 2012) and reinforcing the view of the automaticity of these processes (see also Casaponsa and Duñabeitia, in press).

Following this view, these results suggest that language switch cost effects in the N250 and N400 components for L2-marked primes are mainly guided by an automatic sub-lexical analysis in our sample of bilinguals. As suggested in the recent extension of the BIA+ model (Van Kesteren et al., 2012), the involvement of the sub-lexical route to determine language membership information is critical. When bilingual participants perform a task in their L1, the language-dependent sub-lexical specificities of unconsciously perceived L2 primes that violate the orthographic regularities of the target language give rise to early incongruity effects due to the saliency of these mismatching representations. The processing of marked L2 prime words is then impoverished, due to the effort needed to map these incongruent L2 sub-lexical representations onto L1 lexical forms. The high similarity between the results associated with L1-incongruent Basque marked masked primes from the bilingual and monolingual samples are in line with this hypothesis. The orthotactic specificity of these strings would make monolinguals process them as non-words lacking a referent in their lexicon. In a parallel manner, bilinguals would initially detect the differential orthographic regularities of these items, and even if the strings are represented in their lexicon(s), their processing would get bogged down at sub-lexical levels (i.e., language-selective lexical access for orthographically marked items; see Casaponsa and Duñabeitia, in press, for evidence supporting this hypothesis). These findings also fit with the predictions of the BIA+ extended model, which suggests that the presence of orthographically salient and distinctive units would immediately activate the language nodes via sub-lexical language information, and these would be read out by the task decision system. In the context of the current experiment in which participants were only aware of the presence of L1 words (namely, the targets), L2 marked masked primes represent a conflict for the task decision system given that their orthotactic regularities do not match those of the target language. This conflict gives rise to the early language-mismatch EEG effects found in the N250 epoch.

In contrast, when sub-lexical cues are not available (i.e., unmarked Basque primes), lexical access is the only key to detect whether the target was preceded by a prime belonging to the same or different language. Considering the existence of a single integrated lexicon and the task demands at hand, bilinguals were expected to show similar co-activation and/or competition effects for orthographically unmarked words from a different language and for words from the same language. This way, the possibility to observe reliable switch cost effects for orthographically unmarked items is highly reduced, given that the mismatch at the lexico-semantic level between these items and their controls in the non-switch conditions is highly similar (i.e., they are equally unrelated to the targets; see Van Wijnendaele and Brysbaert, 2002, for similar results during L1 and L2 masked primes with highly proficient bilinguals). Our results for unmarked L2 items are also in line with the BIA+ extended model. In the absence of sub-lexical information pointing to a specific language, L2 and L1 are granted language-independent lexical access to an integrated lexicon, and language identification occurs at (post-)

lexical stages of processing (see also Casaponsa and Duñabeitia, in press). Thereby, similar effects are expected to emerge for L1 and orthographically unmarked L2 masked primes, as demonstrated in the current study. However, in the case of the monolinguals, unmarked L2 words were pseudo-words in essence (i.e., orthographically unmarked strings lacking lexical representation), and they elicited greater negativities due to the effort needed to map the existent orthographic representations onto the inexistent lexical entry (Grainger and Holcomb, 2009). Thus, targets preceded by primes that do not exist in the lexicon (orthographically unmarked items) give rise to greater conflict than those preceded by primes with existing lexical representations (L1 items).

In sum, this study is illustrative of how bilinguals process different languages during reading and of how multilingual reading is grounded in language selection criteria that are based on both basic-level orthographic regularities and on higher-level lexical dimensions. As evidenced by the early effects found only for the marked L2 words, it is suggested that masked language switch priming effects are highly sensitive to the language-dependent sub-lexical features of words. Furthermore, unconsciously perceived L2 primes that matched the orthotactic regularities of the L1 yielded similar conflict to that produced by L1 primes (always unrelated to the target), evidencing the nature of language-independent lexical access in highly proficient bilinguals. These data support the recent extension of the BIA+ model proposed by Van Kesteren et al. (2012) and recent results showing the relevance of sub-lexical cues in automatic and unconscious language-selective and language-nonselective activation by Casaponsa and Duñabeitia (in press). We conclude that the bilingual visual word recognition system is strongly influenced by sub-lexical stages of processing, and that language-dependent orthotactic combinatorial rules play an important role in bilingual lexical access and language identification.

4. Materials and Methods

4.1 Experiment 1

4.1.1. Participants

Twenty-one right-handed native Spanish speakers (10 women; mean age=21, SD= 2.07) completed this experiment as part of the bilingual sample. All of them were recruited from the bilingual pool of participants from the Basque Country. Spanish was their dominant language, but they were also highly proficient in Basque as calculated by self-ratings (see Table 3), and acquired Basque very early in life (mean=2.48, SD=1.66). Their overall self-rated proficiency in Spanish (L1) ranged from 9 to 10 (mean= 9.73, SD=0.46) and their overall Basque (L2) proficiency ranged from 7 to 8 (mean=7.56, SD=0.60). Further self-ratings regarding speaking, understanding, writing and reading skills showed a significant difference in the relative proficiency between languages (i.e., Spanish-dominant; ps<.001). All participants had normal or corrected-to-normal vision and participated voluntarily in this experiment in exchange for monetary compensation. None of the participants reported neurological or psychiatric disorders. Written informed consent was obtained from all participants prior to the experimental session in accordance with guidelines approved by the Ethics and Research Committees of the Basque Center on Cognition, Brain and Language. The study was also performed in accordance with the ethical standards set in the Declaration of Helsinki.

-Insert Table 3 around here-

4.1.2. Stimuli

Three hundred and forty Spanish (L1) words were selected as targets (e.g., cuento [story]), taken from the Spanish B-Pal database (Davis and Perea, 2005) and were split into two groups balanced for word frequency, length, number of neighbors, AoA and concreteness scores (see Table 4). Primes were either unrelated Basque (L2) words (340 items) or unrelated Spanish (L1) words (340 items) selected from the E-Hitz database (Perea et al., 2006) and B-Pal, respectively.

Half of the Spanish targets were preceded by 1) an orthographically marked unrelated Basque word containing implausible bigrams when measured in the L1 (e.g., neska [girl]; "sk" is an implausible bigram in Spanish), or by 2) a Spanish unrelated word (e.g., bolsa [bag]). The other half of Spanish target words could be preceded by 1) an orthographically unmarked unrelated Basque word containing only plausible Spanish bigrams (e.g., mutil [boy]), or by 2) a Spanish unrelated word (e.g., cerebro [brain]). See Table 4 for means and SD of the factors balanced within and across languages.

To control for the L1 plausibility of the L2 primes (i.e., their orthographic markedness), we first obtained all the possible Spanish bigrams from the Spanish LEXESP (Sebastián-Gallés et al., 2000) and Basque SYLLABARIUM (Dunabeitia et al., 2010a) databases respectively, and calculated their log10 frequencies (see the B-Pal database for a similar procedure). These Spanish bigram frequencies were used to split Basque words into two sets, depending on whether their mean bigram frequencies when measured according to the Spanish bigram distribution resulted above or below the mean score. The unmarked Basque primes exclusively included words whose mean bigram frequency fell above the mean of this distribution of Spanish bigrams. Similarly, the marked Basque set only included words whose mean bigram frequency fell below the mean. Furthermore, in order to better capture the differences between sets, another restriction was imposed: All the Basque words from the unmarked set included bigrams that appeared at least 10 times in the Spanish lexicon and all the Basque words from the marked set included at least one bigram that appeared less than 10 times in the Spanish lexicon. This double-check at the individual bigram frequency level and at the mean bigram frequency level provides a reliable measure to assess the bigram probability of appearance or plausibility of Basque words in Spanish and also allowed us to match Basque and Spanish word sets in terms of their bigram frequencies within and across languages (see Table 4). Critically, the unmarked primes were as frequent when measured according to the Spanish bigram count as the words from the Spanish sets. Therefore, the unmarked Basque words were orthographically very similar to the Spanish vocabulary.

Finally, prime-target orthographic overlap was also balanced across conditions by matching the Levenshtein distance scores of all the possible prime-target combinations (Mean=8.38, SD=1.86; all ps>.71; Minimum number of Levenshtein edits = 5). In addition, two lists were created and counterbalanced across participants, so that each target word appeared only once in each list but in a different prime condition in each of them. Participants were randomly assigned to each list and priming conditions were evenly distributed across and within lists (340 critical prime-target pair words in each list, with 85 words pairs per condition in each list). All these pairs were used as no-go trials in a semantic categorization task. An additional set of 60 Spanish animal names were also included (15% of the final trials) for the go-trials and for the prime visibility test in each of the lists. The prime visibility test consisted of the inclusion of the animal names as masked primes, followed by the presentation of non-animal targets, and it is typically used in EEG masked priming studies in order to make sure that participants are not aware of the presence and nature of the masked primes (see, among many others, Molinaro, Duñabeitia, Marín-Gutiérrez, and Carreiras, 2010, for a similar procedure). (Note that if participants respond to these masked animal names, it would suggest that they consciously perceive the briefly presented primes). The complete list of materials used in this experiment can be accessed through http://www.bcbl.eu/materials/acasaponsa/Appendix_Materials_C&C&D.pdf

-Insert Table 4 around here-

4.1.3. Procedure

All participants were tested individually in a quiet room. Visual stimuli were presented using Presentation software (Version 4.6, Neurobehavioral systems, Inc.) on a 15” CRT monitor set to a refresh rate of 90Hz. Stimuli were displayed at high contrast in white letters on a black background. On each trial, a forward mask consisting of a row of hash marks (#) was presented for 500ms. Next, the prime was presented in 25pt lowercase Courier New and stayed on the screen for 50 ms. The prime was immediately followed by the presentation of the target stimulus in 25pt uppercase Courier New. The target remained on the screen for 500 ms. The inter-trial interval varied randomly between 900 and 1100 ms. After this interval, an asterisk was presented for 1000 ms in order to allow for participant blinks. Participants were instructed to press the space bar in the keyboard whenever they saw an animal name on the screen and to read all other words passively without responding to them. The fact that critical stimuli did not require an over response (i.e., semantic categorization go/no go task) granted a ERPs signal free from muscular artifacts. Participants were not informed of the presence of the primes. Trial presentation order was randomized across participants. Each participant received a total of 20 practice trials (representative of the conditions in the critical trials) prior the experimental trials. Task instructions (and interactions with the participants) were given in their L1 (Spanish). The experimental session lasted for approximately 20 minutes (excluding participant preparation).

4.1.4. EEG recording procedure

The electroencephalogram was recorded from 27 active electrodes (plus Ground) held in place on the scalp by an elastic cap (ElectroCap International, Eaton, USA, 10-10 system). Eye movements and blinks were monitored with four further electrodes providing bipolar recordings of the horizontal (Heog-, Heog+) and vertical (Veog-, Veog+) electrooculogram (EOG). Another two electrodes were attached over the right mastoid bone (reference) and over the left mastoid bone (recorded actively to monitor for differential mastoid activity). All EEG electrode impedances were maintained below 5 kΩ (impedance for eye electrodes was less than 10 kΩ).

The EEG signal was sampled continuously throughout the experiment with a sampling rate of 500 Hz and digitally off-line re-referenced to linked mastoids.

4.1.5. Data analysis

Ocular artifacts were corrected using Independent Component Analysis (ICA). Based on previous literature, the ICA algorithm used was Infomax (Gradient) Restricted Biased. High-pass filter of 0.01 Hz was applied before the ICA procedure, and a low-pass filter of 30 Hz was applied after ICA. Averaged ERPs time-locked to target onset were formed off-line from trials free of ocular and muscular artifacts (85.27% of the data; rejected trials were equally distributed across conditions). Baseline correction was performed using the averaged EEG activity in the 100 ms preceding the onset of the target stimuli. ERPs were quantified by taking the mean amplitude of each participant and electrode in two temporal epochs corresponding to the two primary components of interest found to be sensitive to masked language switch cost ERP effects: the N250 (covering a time window between 180 and 260 ms post-target onset) and the N400 (represented by a time window between 350 and 500 ms). These two epochs of interest are in accordance with those reported in earlier studies on a similar topic using a parallel procedure (e.g., Chauncey et al., 2008, 2011; Dunabeitia et al., 2010b; Holcomb and Grainger, 2006; Hoshino et al., 2010; Midgley et al., 2009b). These two epochs were analyzed separately in two sets of repeated measures analyses of variance (ANOVAs). 21 out of the 27 active electrodes were used for the analysis, creating the factor Region with three different levels formed by the averaged amplitude of 7 adjacent electrodes: Anterior (Fp1|Fp2|F7|F3|Fz|F4|F8), Central (FC5|FC1|FC2|FC6|C3|Cz|C4|) and Posterior (CP5|CP1|CP2|CP6|P3|Pz|P4). Together with the Region factor, the two other factors associated with the design were included in the analyses: Language (2 levels: switch|non-switch) and Bigram (2 levels: marked|unmarked). Greenhouse-Geisser correction was applied for departure from sphericity (Greenhouse and Geisser, 1959).

Greenhouse-Geisser epsilon value (ε) is provided only when different from 1, informing that there was a violation of the assumption of sphericity, and the corrected *p*-value is therefore reported. Effect-size was estimated using the partial eta-squared coefficient η^2_p (Cohen, 1973; Haase, 1983).

4.2. Experiment 2

4.2.1. Participants

Twenty-one right-handed native Spanish speakers (15 women; mean age=20.86, SD= 3.23) with normal or corrected-to-normal vision participated voluntarily in the experiment as the monolingual sample. All of them were recruited from a pool of monolingual participants from the University of Murcia. In contrast to the Basque Country, where both Spanish and Basque are the official languages, Murcia is an autonomous community of Spain where Spanish is the only official language. None of the participants had any knowledge of Basque at all. Their overall self-rated proficiency in Spanish ranged from 8 to 10 (mean= 9.48, SD=0.75). All participants had normal or corrected-to-normal vision and participated voluntarily in this experiment in exchange for course credits. None of the participants reported neurological or psychiatric disorders. Written informed consent was obtained from all participants prior to the experimental session in accordance with guidelines approved by the Ethics and Research Committees of the Basque Center on Cognition, Brain and Language. The study was also performed in accordance with the ethical standards set in the Declaration of Helsinki.

4.2.3. Materials, Procedure and Data Analysis. These were the same as in Experiment 1.

Acknowledgements

This research was partially supported by the Spanish Government (CSD2008-00048, PSI2012-32123 and PSI2012-31448), the European Research Council (ERC-AdG-295362) and the Basque Government (PI2012-74). The authors are grateful to Guillaume Thierry, Margaret Gillon Dowens, Ileana Quiñones and Ainhoa Bastarrika for their helpful comments on earlier drafts and for stimulating discussion on the interpretation of the results.

References

Carreiras, M., Duñabeitia, J. A. & Molinaro, N. (2009a). Consonants and vowels contribute differently to visual word recognition: ERPs of relative position priming. Cerebral Cortex, 19, 2659 – 2670. doi: 10.1093/cercor/bhp019

Carreiras, M., Gillon-Dowens, M., Vergara, M. & Perea, M. (2009b). Are vowels and consonants processed differently? ERP evidence with a delayed letter paradigm. Journal of Cognitive Neuroscience. 21, 275-288. doi: 10.1162/jocn.2008.21023

Casaponsa, A., Carreiras, M., & Dunabeitia, J. A. (2014). Discriminating languages in bilingual contexts: the impact of orthographic markedness. *Front Psychol, 5*, 424. doi: 10.3389/fpsyg.2014.00424

Casaponsa, A., & Duñabeitia, J. A (in press). Lexical organization of language-ambiguous and language-specific words in bilinguals. Quarterly Journal of Experimental Psychology.

Chauncey, K., Grainger, J., & Holcomb, P. J. (2008). Code-switching effects in bilingual word recognition: a masked priming study with event-related potentials. *Brain Lang, 105*(3), 161-174. doi: 10.1016/j.bandl.2007.11.006

Chauncey, K., Grainger, J., & Holcomb, P. J. (2011). The role of subjective frequency in language switching: an ERP investigation using masked priming. *Mem Cognit, 39*(2), 291-303. doi: 10.3758/s13421-010-0006-7

Coch, D., & Mitra, P. (2010). Word and pseudoword superiority effects reflected in the ERP waveform. *Brain Res, 1329*, 159-174. doi: 10.1016/j.brainres.2010.02.084

Cohen, J. (1973). Eta-squared and partial eta-squared in fixed factor ANOVA designs. *Educational and psychological measurement*.

Dandurand, F., Grainger, J., Dunabeitia, J. A., & Granier, J. P. (2011). On coding non-contiguous letter combinations. *Front Psychol, 2*, 136. doi: 10.3389/fpsyg.2011.00136

Davis, C. J., & Perea, M. (2005). BuscaPalabras: A program for deriving orthographic and phonological neighborhood statistics and other psycholinguistic indices in Spanish. *Behavior Research Methods, 37*(4), 665-671.

Dijkstra, T., Grainger, J., & Van Heuven, W. J. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language, 41*(4), 496-518.

Dijkstra, T., Timmermans, M., & Schriefers, H. (2000). On being blinded by your other language: Effects of task demands on interlingual homograph recognition. *Journal of Memory and Language, 42*(4), 445-464.

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*(03). doi: 10.1017/s1366728902003012

Dimitropoulou, M., Dunabeitia, J. A., & Carreiras, M. (2011a). Masked translation priming effects with low proficient bilinguals. *Mem Cognit, 39*(2), 260-275. doi: 10.3758/s13421-010-0004-9

Dimitropoulou, M., Dunabeitia, J. A., & Carreiras, M. (2011b). Two words, one meaning: evidence of automatic co-activation of translation equivalents. *Front Psychol, 2*, 188. doi: 10.3389/fpsyg.2011.00188

Dunabeitia, J. A., Cholin, J., Corral, J., Perea, M., & Carreiras, M. (2010b). SYLLABARIUM: an online application for deriving complete statistics for Basque and Spanish orthographic syllables. *Behav Res Methods, 42*(1), 118-125. doi: 10.3758/BRM.42.1.118

Duñabeitia, J. A., Dimitropoulou, M., Uribe-Etxebarria, O., Laka, I., & Carreiras, M. (2010a). Electrophysiological correlates of the masked translation priming effect with highly proficient simultaneous bilinguals. *Brain Res, 1359*, 142-154. doi: 10.1016/j.brainres.2010.08.066

Duñabeitia, J. A., Molinaro, N., Laka, I., Estévez, A., & Carreiras, M. (2009). N250 effects for letter transpositions depend on lexicality: 'casual' or 'causal'? *NeuroReport, 20*(4), 381-387. doi: 10.1097/WNR.0b013e3283249b1c

Duyck, W., Assche, E. V., Drieghe, D., & Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: evidence for nonselective lexical access. *J Exp Psychol Learn Mem Cogn, 33*(4), 663-679. doi: 10.1037/0278-7393.33.4.663

Duyck, W., & Warlop, N. (2009). Translation priming between the native language and a second language: new evidence from Dutch-French bilinguals. *Exp Psychol, 56*(3), 173-179. doi: 10.1027/1618-3169.56.3.173

Grainger, J., & Beauvillain, C. (1987). Language blocking and lexical access in bilinguals. *The Quarterly Journal of Experimental Psychology, 39*(2), 295-319.

Grainger, J., & Holcomb, P. J. (2009). Watching the Word Go by: On the Time-course of Component Processes in Visual Word Recognition. *Lang Linguist Compass, 3*(1), 128-156. doi: 10.1111/j.1749-818X.2008.00121.x

Grainger, J., & Van Heuven, W. J. (2003). Modeling letter position coding in printed word perception. *Ment Lex*, 1-24.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24*(2), 95-112.

Grossi, G., & Coch, D. (2005). Automatic word form processing in masked priming: an ERP study. *Psychophysiology, 42*(3), 343-355. doi: 10.1111/j.1469-8986.2005.00286.x

Grossi, G., Savill, N., Thomas, E., & Thierry, G. (2012). Electrophysiological cross-language neighborhood density effects in late and early english-welsh bilinguals. *Front Psychol, 3*, 408. doi: 10.3389/fpsyg.2012.00408

Haase, R. F. (1983). Classical and partial eta square in multifactor ANOVA designs. *Educational and psychological measurement, 43*(1), 35-39.

Hauk, O., Davis, M. H., Ford, M., Pulvermuller, F., & Marslen-Wilson, W. D. (2006). The time course of visual word recognition as revealed by linear regression analysis of ERP data. *NeuroImage, 30*(4), 1383-1400. doi: 10.1016/j.neuroimage.2005.11.048

Hauk, O., Davis, M. H., & Pulvermuller, F. (2008). Modulation of brain activity by multiple lexical and word form variables in visual word recognition: A parametric fMRI study. *NeuroImage, 42*(3), 1185-1195. doi: 10.1016/j.neuroimage.2008.05.054

Holcomb, P. J., & Grainger, J. (2006). On the time course of visual word recognition: an event-related potential investigation using masked repetition priming. *J Cogn Neurosci, 18*(10), 1631-1643. doi: 10.1162/jocn.2006.18.10.1631

Hoshino, N., Midgley, K. J., Holcomb, P. J., & Grainger, J. (2010). An ERP investigation of masked cross-script translation priming. *Brain Res, 1344*, 159-172. doi: 10.1016/j.brainres.2010.05.005

Hoshino, N., & Thierry, G. (2012). Do Spanish-English Bilinguals have Their Fingers in Two Pies - or is It Their Toes? An Electrophysiological Investigation of Semantic Access in Bilinguals. *Front Psychol, 3*, 9. doi: 10.3389/fpsyg.2012.00009

Lemhöfer, K., & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition, 32*(4), 533-550.

Lemhöfer, K., Dijkstra, T., Schriefers, H., Baayen, R. H., Grainger, J., & Zwitserlood, P. (2008). Native language influences on word recognition in a second language: a megastudy. *J Exp Psychol Learn Mem Cogn, 34*(1), 12-31. doi: 10.1037/0278-7393.34.1.12

Lemhöfer, K., Koester, D., & Schreuder, R. (2011). When bicycle pump is harder to read than bicycle bell: effects of parsing cues in first and second language compound reading. *Psychon Bull Rev, 18*(2), 364-370. doi: 10.3758/s13423-010-0044-y

Martin, C. D., Costa, A., Dering, B., Hoshino, N., Wu, Y. J., & Thierry, G. (2012). Effects of speed of word processing on semantic access: the case of bilingualism. *Brain Lang, 120*(1), 61-65. doi: 10.1016/j.bandl.2011.10.003

Massol, S., Midgley, K. J., Holcomb, P. J., & Grainger, J. (2011). When less is more: feedback, priming, and the pseudoword superiority effect. *Brain Res, 1386*, 153-164. doi: 10.1016/j.brainres.2011.02.050

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009a). Language effects in second language learners and proficient bilinguals investigated with event-related potentials. *J Neurolinguistics, 22*(3), 281-300. doi: 10.1016/j.jneuroling.2008.08.001

Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009b). Masked repetition and translation priming in second language learners: a window on the time-course of form and meaning activation using erps. *Psychophysiology, 46*(3), 551-565. doi: 10.1111/j.1469-8986.2009.00784.x

Midgley, K. J., Holcomb, P. J., Van Heuven, W. J., & Grainger, J. (2008). An electrophysiological investigation of cross-language effects of orthographic neighborhood. *Brain Res, 1246*, 123-135. doi: 10.1016/j.brainres.2008.09.078

Molinaro, N., Duñabeitia, J.A., Marín-Gutiérrez, A. & Carreiras, M. (2010). From numbers to letters: Feedback regularization in visual word recognition. *Neuropsychologia, 48* (5), 1343-1355. doi: 10.1016/j.neuropsychologia.2009.12.037

Morris, J., Frank, T., Grainger, J., & Holcomb, P. J. (2007). Semantic transparency and masked morphological priming: an ERP investigation. *Psychophysiology, 44*(4), 506-521. doi: 10.1111/j.1469-8986.2007.00538.x

Ng, S., & Wicha, N. Y. (2013). Meaning first: a case for language-independent access to word meaning in the bilingual brain. *Neuropsychologia, 51*(5), 850-863. doi: 10.1016/j.neuropsychologia.2013.01.017

Orfanidou, E., & Sumner, P. (2005). Language switching and the effects of orthographic specificity and response repetition. *Memory & Cognition, 33*(2), 355-369.

Perea, M., Duñabeitia, J. A., & Carreiras, M. (2008). Masked associative/semantic priming effects across languages with highly proficient bilinguals. *Journal of Memory and Language, 58*(4), 916-930. doi: 10.1016/j.jml.2008.01.003

Perea, M., Urkia, M., Davis, C. J., Agirre, A., Laseka, E., & Carreiras, M. (2006). E-Hitz: a word frequency list and a program for deriving psycholinguistic statistics in an agglutinative language (Basque). *Behav Res Methods, 38*(4), 610-615.

Proverbio, A. M., Adorni, R., & Zani, A. (2009). Inferring native language from early bio-electrical activity. *Biol Psychol, 80*(1), 52-63.

Proverbio, A. M., Leoni, G., & Zani, A. (2004). Language switching mechanisms in simultaneous interpreters: an ERP study. *Neuropsychologia, 42*(12), 1636-1656.

Schwartz, A. I., Kroll, J. F., & Diaz, M. (2007). Reading words in Spanish and English: Mapping orthography to phonology in two languages. *Language and Cognitive Processes, 22*(1), 106-129. doi: 10.1080/01690960500463920

Sebastián-Gallés, N., Martí, M., Carreiras, M., & Cuetos, F. (2000). LEXESP: Una base de datos informatizada del español. *Universitat de Barcelona, Barcelona*.

Spalek, K., Hoshino, N., Wu, Y. J., Damian, M., & Thierry, G. (2014). Speaking two languages at once: Unconscious native word form access in second language production. *Cognition, 133*(1), 226-231.

Thierry, G., & Wu, Y. J. (2004). Electrophysiological evidence for language interference in late bilinguals. *NeuroReport, 15*(10), 1555-1558. doi: 10.1097/01.wnr.0000134214.57469.c2

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proc Natl Acad Sci U S A, 104*(30), 12530-12535. doi: 10.1073/pnas.0609927104

Thomas, M. S. C., & Allport, A. (2000). Language Switching Costs in Bilingual Visual Word Recognition. *Journal of Memory and Language, 43*(1), 44-66. doi: 10.1006/jmla.1999.2700

Vaid, J., & Frenck-Mestre, C. (2002). Do orthographic cues aid language recognition? A laterality study with French–English bilinguals. *Brain and Language, 82*(1), 47-53.

Van Der Meij, M., Cuetos, F., Carreiras, M., & Barber, H. A. (2011). Electrophysiological correlates of language switching in second language learners. *Psychophysiology, 48*(1), 44-54. doi: 10.1111/j.1469-8986.2010.01039.x

Van Heuven, W. J., Dijkstra, T., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory and Language, 39*(3), 458-483.

Van Heuven, W. J., Schriefers, H., Dijkstra, T., & Hagoort, P. (2008). Language conflict in the bilingual brain. *Cerebral Cortex, 18*(11), 2706-2716.

Van Kesteren, R., Dijkstra, T., & de Smedt, K. (2012). Markedness effects in Norwegian-English bilinguals: task-dependent use of language-specific letters and bigrams. *Q J Exp Psychol (Hove), 65*(11), 2129-2154. doi: 10.1080/17470218.2012.679946

Van Wijnendaele, I., & Brysbaert, M. (2002). Visual word recognition in bilinguals: phonological priming from the second to the first language. *Journal of Experimental Psychology: Human Perception and Performance, 28*(3), 616.

Von Studnitz, R. E., & Green, D. W. (2002). Interlingual homograph interference in German–English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition, 5*(01), 1-23.

Whitney, C., Bertrand, D., & Grainger, J. (2011). On coding the position of letters in words: a test of two models. *Exp Psychol, 59*(2), 109-114. doi: 10.1027/1618-3169/a000132

Whitney, C., & Cornelissen, P. (2008). SERIOL Reading. *Language and Cognitive Processes, 23*(1), 143-164. doi: 10.1080/01690960701579771

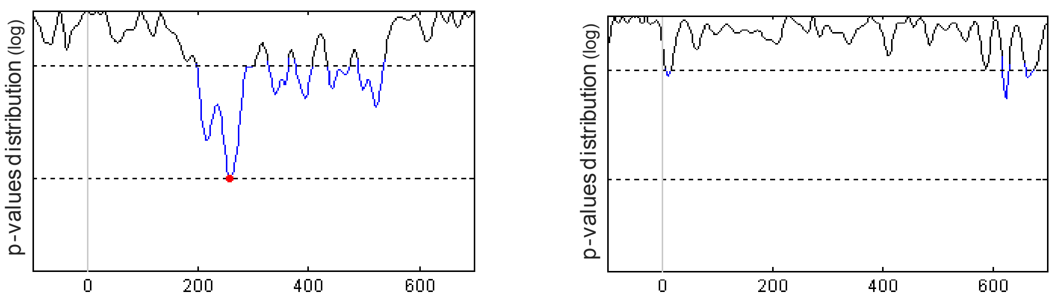
Zhang, T., Van Heuven, W. J., & Conklin, K. (2011). Fast automatic translation and morphological decomposition in Chinese-English bilinguals. *Psychol Sci, 22*(10), 1237-1242. doi: 10.1177/0956797611421492

Appendix

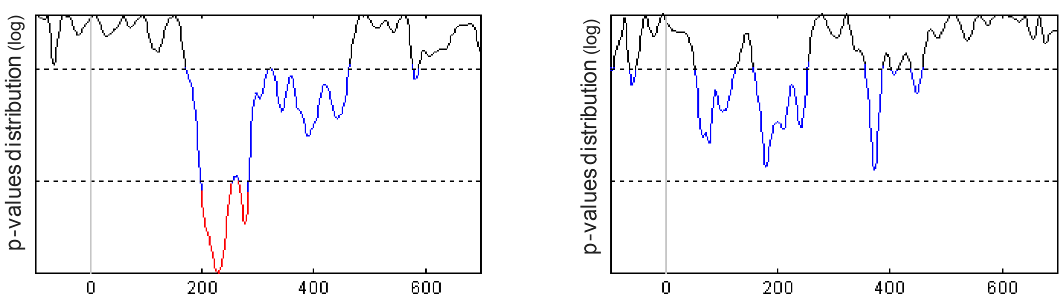
Distribution of the masked language effects across time.

A) Bilingual group

Marked effect | **Unmarked effect**



B) Monolingual group



Note: p-values distribution across the trial for each time-point collapsed across all electrodes and transformed to a logarithmic scale for visualization purposes. Blue color reflects p-values above uncorrected significance level (α=0.05). Red color reflects p-values above the level of significance corrected for multiple comparisons using Bonferroni adjustment (see Dunn, 1961). The value of α-adjusted was 0.000125, corresponding to the 400 comparisons, one for each time-point (i.e., 2 ms).

Figure captions

Figure 1. Experiment 1: A) ERPs associated with the language switch (thick lines) and non-switch conditions (thin lines) in the marked (upper) and unmarked (lower) bigram sets. B) Topographical distribution of the masked language switch cost effects in terms of amplitude differences between the unrelated Basque primes and unrelated Spanish primes. Differences are plotted separately for the averaged activity between 180 and 260 ms (N250) and 350 to 500 ms (N400).

Figure 2. Experiment 2: A) ERPs associated with the language switch (thick lines) and non-switch conditions (thin lines) in the marked (upper) and unmarked (lower) bigram sets. B) Topographical distribution of the masked language switch cost effects in terms of amplitude differences between the unrelated Basque primes and unrelated Spanish primes. Differences are plotted separately for the averaged activity between 180 and 260 ms (N250) and 350 to 500 ms (N400).

Table 1. Voltatge means and standard deviations for each condition and time window used in Experiment 1.

	Marked						Unmarked					
	Spanish			Basque			Spanish			Basque		
N250	Anterior	Central	Posterior	Anterior	Central	Posterior	Anterior	Central	Posterior	Anterior	Central	Posterior
Mean	3.98	2.44	1.26	2.95	1.21	.02	3.53	1.96	.96	3.43	1.82	.57
SD	2.56	2.54	3.04	2.74	2.41	2.82	2.62	2.62	3.11	3.01	2.94	3.24
N400												
Mean	.42	.15	.95	-.13	-.40	.48	-.12	-.50	.39	.04	-.45	.34
SD	2.37	2.31	2.22	2.35	2.30	2.56	2.75	2.78	2.55	2.73	2.81	2.71

Table 2. Voltatge means and standard deviations for each condition and time window used in Experiment 2.

	Marked						Unmarked					
	Spanish			Basque			Spanish			Basque		
N250	Anterior	Central	Posterior	Anterior	Central	Posterior	Anterior	Central	Posterior	Anterior	Central	Posterior
Mean	3.89	2.93	1.49	2.61	1.25	-.02	3.94	2.78	1.39	3.35	1.97	.54
SD	3.36	3.28	2.36	2.67	2.74	2.47	2.81	2.66	2.13	2.87	2.72	2.18
N400												
Mean	-.22	.19	.46	-.91	-.59	.07	-.45	-.10	.38	-.87	-.72	-.31
SD	2.31	2.25	2.07	2.18	2.12	1.98	2.33	2.31	1.91	1.94	1.58	1.75

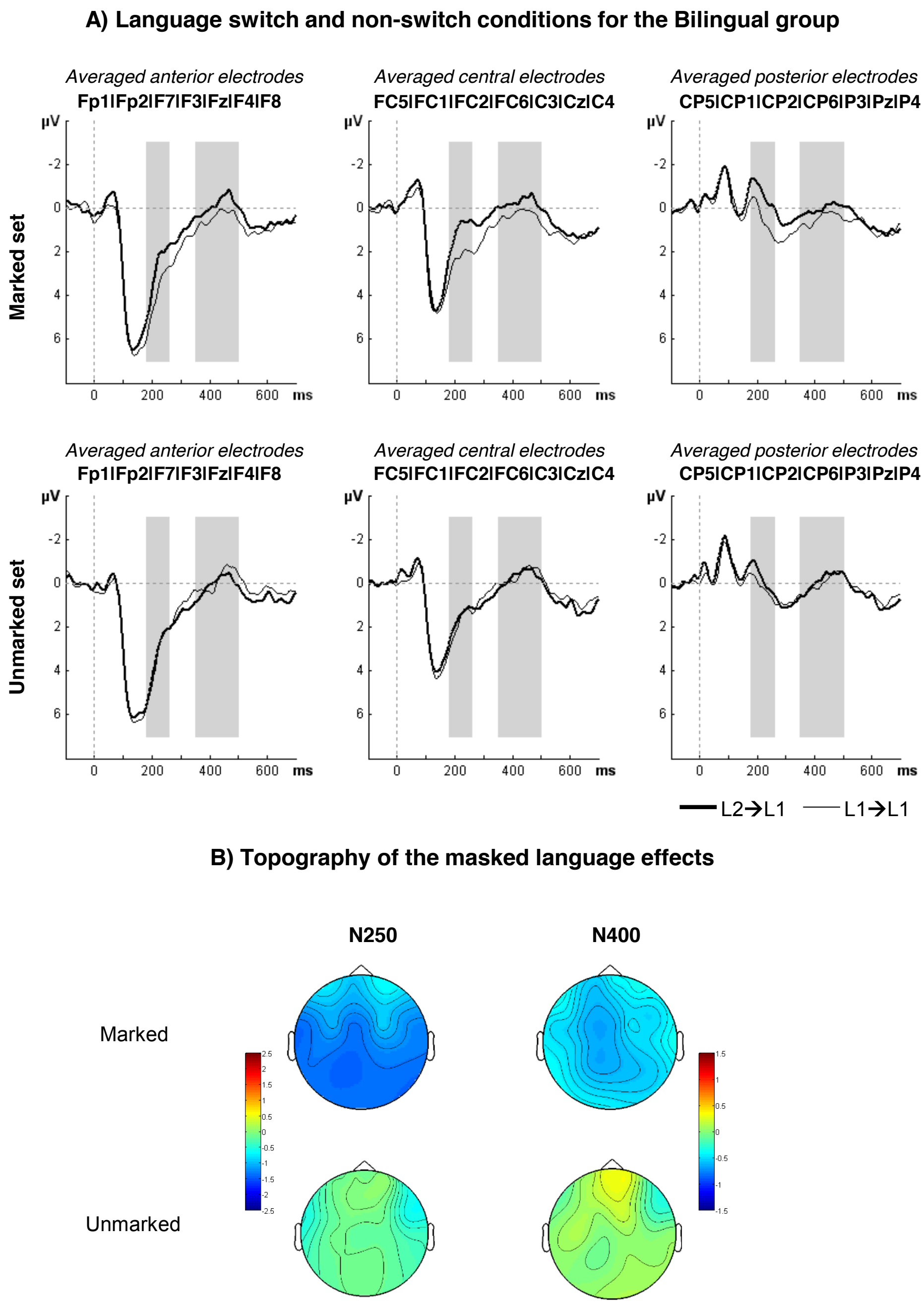
Table 3. Mean levels of L1 (Spanish) and L2 (Basque) language proficiency calculated according to participants' self-ratings (in a 1-to-10 scale) for Experiment 1 and 2. Standard deviations are provided within parentheses. All between-language comparisons resulted in ps<0.001.

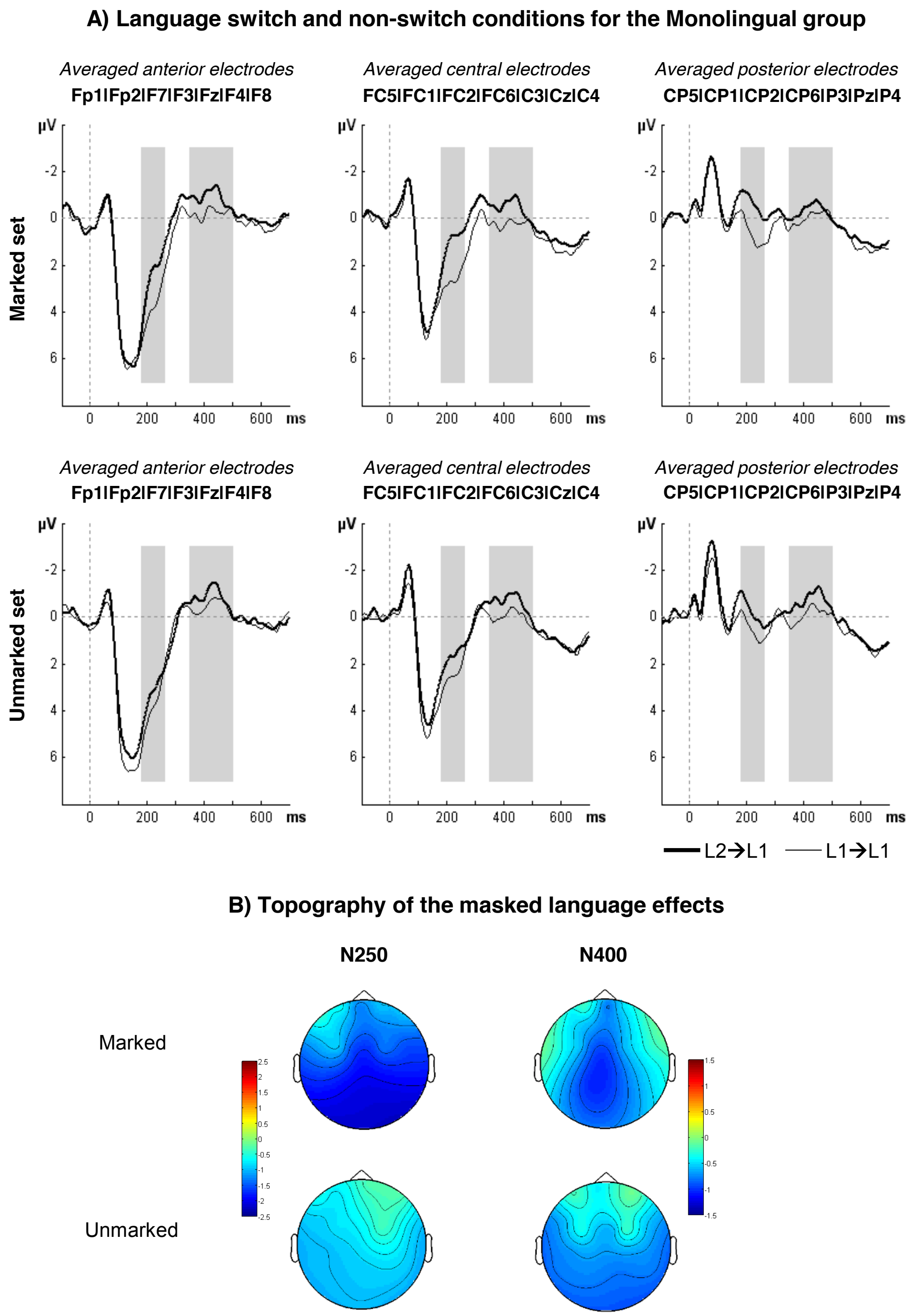
	Language proficiency			
	Experiment 1 (Bilinguals)		Experiment 2 (Monolinguals)	
	Spanish	Basque	Spanish	Basque
Speaking	9.91 (0.29)	7.22 (0.97)	9.67 (0.58)	-
Understanding	9.86 (0.35)	8.45 (1.26)	9.54 (0.74)	-
Writing	9.81 (0.39)	7.55 (1.14)	9.57 (0.68)	-
Reading	9.86 (0.35)	8.55 (1.14)	9.55 (0.84)	-
General self-perception	9.73 (0.46)	7.55 (0.60)	9.48 (0.75)	-

Table 4. Mean values for each sub-lexical, lexical and semantic factor of the L1 (Spanish) and L2 (Basque) word used in Experiment 1 and 2 split by condition. Standard deviations are provided within parentheses. Within-language and across language p-values are also provided.

	PRIME WORDS								TARGET WORDS		
	BASQUE			SPANISH			Between-language Comparisons (p-values)		SPANISH		
	Marked	Unmarked	p-values	Marked control	Unmarked control	p-values	Marked	Unmarked	Marked	Unmarked	p-values
Word Frequency	52.00 (114.53)	47.36 (109.53)	0.70	44.65 (81.17)	42.56 (74.86)	0.81	0.50	0.64	43.07 (62.05)	38.47 (48.98)	0.45
Word Length	6.62 (1.83)	6.81 (2.22)	0.35	6.81 (1.81)	6.82 (1.77)	0.98	0.33	0.98	7.65 (2.27)	7.63 (2.11)	0.94
Number of Orthographic Neighbors	1.42 (1.62)	1.55 (0.35)	0.54	1.53 (2.74)	1.69 (3.01)	0.61	0.67	0.64	1.96 (3.60)	1.69 (3.00)	0.45
Age of Acquisition	3.22 (0.49)	3.23 (0.50)	0.82	3.19 (0.56)	3.19 (0.61)	0.98	0.55	0.45	3.17 (0.57)	3.18 (0.60)	0.86
Word Concreteness	4.09 (0.89)	4.12 (0.86)	0.81	4.05 (0.81)	4.07 (0.85)	0.80	0.65	0.65	3.85 (0.82)	3.79 (0.84)	0.53
Spanish Bigram Frequency	1.72 (0.3)	2.97 (0.24)	0.00	2.49 (0.30)	2.46 (0.33)	0.42	0.00	0.34			
Basque Bigram Frequency	2.88 (0.18)	2.89 (0.20)	0.68								
Number of Spanish-Implausible Bigrams	2.35 (0.93)	0 (0)	0.00								

Note: Frequency ratings were obtained from the B-Pal and E-Hitz databases. Bigram frequencies were obtained from the raw number of appearances of the bigrams in the Spanish LEXESP and Basque SYLLABARIUM corpus in a within-word position-independent manner. The critical factor Number of Spanish-Implausible Bigrams refers to the number of bigrams from the Basque words that appear less than 10 times in the whole Spanish corpus.


A) Language switch and non-switch conditions for the Bilingual group
Averaged anterior electrodes
Fp1|Fp2|F7|F3|Fz|F4|F8
Averaged central electrodes
FC5|FC1|FC2|FC6|C3|Cz|C4
Averaged posterior electrodes
CP5|CP1|CP2|CP6|P3|Pz|P4
Marked set
Unmarked set
µV
ms
L2→L1
L1→L1
B) Topography of the masked language effects
N250
N400
Marked
Unmarked


A) Language switch and non-switch conditions for the Monolingual group
Averaged anterior electrodes
Fp1|Fp2|F7|F3|Fz|F4|F8
Averaged central electrodes
FC5|FC1|FC2|FC6|C3|Cz|C4
Averaged posterior electrodes
CP5|CP1|CP2|CP6|P3|Pz|P4
Marked set
Unmarked set
µV
ms
L2→L1
L1→L1
B) Topography of the masked language effects
N250
N400
Marked
Unmarked