BRAIN RESEARCH 1542 (2014) 111–130

Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

Research Report

Head start for target language in bilingual listening

Ian FitzPatrick[a,b,c,*], Peter Indefrey[a,b,c]

[a]Heinrich Heine University Düsseldorf, Germany
[b]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[c]Radboud University Nijmegen, Donders Institute for Brain, Cognition, and Behaviour, Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Accepted 10 October 2013
Available online 18 October 2013

Keywords:
Language
Bilingualism
EEG
Event-related-potentials
N400
LPC

ABSTRACT

In this study we investigated the availability of non-target language semantic features in bilingual speech processing. We recorded EEG from Dutch–English bilinguals who listened to spoken sentences in their L2 (English) or L1 (Dutch). In Experiments 1 and 3 the sentences contained an interlingual homophone. The sentence context was either biased towards the target language meaning of the homophone (*target biased*), the non-target language meaning (*non-target biased*), or neither meaning of the homophone (*fully incongruent*). These conditions were each compared to a semantically congruent control condition. In L2 sentences we observed an N400 in the *non-target biased* condition that had an earlier offset than the N400 to *fully incongruent* homophones. In the *target biased* condition, a negativity emerged that was later than the N400 to *fully incongruent* homophones. In L1 contexts, neither *target biased* nor *non-target biased* homophones yielded significant N400 effects (compared to the control condition). In Experiments 2 and 4 the sentences contained a language switch to a non-target language word that could be semantically congruent or incongruent. Semantically incongruent words (switched, and non-switched) elicited an N400 effect. The N400 to semantically congruent language-switched words had an earlier offset than the N400 to incongruent words. Both congruent and incongruent language switches elicited a Late Positive Component (LPC). These findings show that bilinguals activate both meanings of interlingual homophones irrespective of their contextual fit. In L2 contexts, the target-language meaning of the homophone has a head start over the non-target language meaning. The target-language head start is also evident for language switches from both L2-to-L1 and L1-to-L2.

1. Introduction

Spoken language is fraught with ambiguity. Words with different meanings often sound identical (e.g., *mail* and *male*) leaving the listener with the task to determine which meaning ('post' or 'man') is contextually appropriate. The ambiguity problem is compounded in bilingual listeners who may additionally be faced with words that sound the same between their languages. The Dutch word *meel* ('flour'), for instance, sounds similar to the English word *mail*. Although words of two different languages are seldom phonologically identical, initial investigations of bilingual word comprehension

*Corresponding author at: Institut für Sprache und Information, Heinrich Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany. Fax: +31 243610989.
E-mail addresses: ian.fitzpatrick@uni-duesseldorf.de, ian@ianfitzpatrick.eu (I. FitzPatrick).

0006-8993/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.brainres.2013.10.014

nevertheless showed concurrent activation of first (L1) and second (L2) language words in bilinguals (Marian and Spivey, 2003a, 2003b; Marian et al., 2003; Spivey and Marian, 1999; Weber and Cutler, 2004). Whether or not this is a problem for bilingual language comprehension depends on the extent to which words from the non-target language (i.e., language not-in-use) compete for recognition and at which point in time the target language (i.e., context language) word emerges as the winner. In the present study we set out to investigate the dynamics of between language word activation in first- and second language speech comprehension.

1.1. *Within-language ambiguity*

Comprehension of ambiguous words has received much attention in studies of monolingual reading and listening. Models of word ambiguity resolution generally fall into one of three categories (for a review see: Simpson and Kang, 1994): (1) context dependent models that assume that only the contextually appropriate meaning of the ambiguous word is accessed at any given moment (e.g., Glucksberg et al., 1986; Simpson, 1981), (2) exhaustive access models that assume that every meaning of the ambiguous word is activated irrespective of the context (Swinney, 1979), or (3) ordered access models that assume that meanings are accessed sequentially from most to least frequent (e.g., Hogaboam and Perfetti, 1975). These models differ greatly in the extent to which they allow for contextual influence on ambiguous word processing. The context is either the driving force behind ambiguity resolution (in the case of context dependent models) or largely unimportant (in the case of exhaustive access and ordered access models). There are also models that allow for influences of both context and word frequency on the access to ambiguous word meanings. The re-ordered access model (Duffy et al., 1988), for instance, assumes that access to ambiguous word meanings proceeds according to word frequency in the absence of a biasing sentence context. However, a strongly biasing context can effectively reorder the sequence of word meaning activations. Thus it seems that sentential bias and the relative frequency of ambiguous word meanings could play a role in determining whether or not multiple meanings of an ambiguous word compete for recognition in monolingual listening. In light of this, two important questions are raised with respect to bilingual language comprehension: is there multiple access to between-language ambiguous word meanings? And if so, to what extent is this cross-linguistic activation modulated by context?

1.2. *Between-language ambiguity*

There is a great deal of evidence that suggests that bilinguals are not capable of restricting their language processing to a single (target) language. The vast majority of this evidence comes from studies of bilingual visual word recognition. Among the most replicated results are findings of a processing advantage for between language cognates (i.e., words that share form and meaning) and a processing cost for interlingual homographs (for an overview see Dijkstra (2005)). These findings have lent weight to models of bilingual language processing that assume that the bilingual language comprehension system is fundamentally language non-selective in nature, such as the Bilingual Interactive Activation (BIA) and BIA+ models (Dijkstra and van Heuven, 1998, 2002; Grainger and Dijkstra, 1992). Beyond visual word recognition there are also reports of cross-linguistic lexical activation in bilingual speech comprehension. Schulpen et al. (2003), for example, showed that both pronunciations of a Dutch–English interlingual near homophone (e.g., /li:f/ - English: *leaf*, Dutch: *lief* 'sweet') could prime the English orthographic form of the word (LEAF). More strikingly, a number of studies that employed the visual world paradigm (Tanenhaus et al., 1995) have shown cross-linguistic lexical activation based solely on word-initial (i.e., incomplete) phonological overlap. Initial phonological overlap of a spoken target word with a visual referent's non-target language translation causes bilinguals to look more often at that referent than at a distractor object with no phonological relation (in either language) to the target (e.g., Spivey and Marian, 1999; Weber and Cutler, 2004). For example, when tasked to "Click on the *desk*" Dutch–English bilinguals tended to fixate on a picture of a lid (Dutch: 'deksel') more often than on a control object such as a swing (Dutch: 'schommel') (FitzPatrick, 2011; Weber and Cutler, 2004).

Curiously, whereas many studies showed robust activation of L1 words from within L2, findings of L2 word access in L1 language processing are more inconsistent (for a review see Degani and Tokowicz (2010)). While Spivey and Marian (Spivey and Marian, 1999; Marian and Spivey, 2003b; see also Lagrou et al. (2011)) show activation of L2 lexical competitors in L1 listening, other studies found L2 activation in L1 processing to be limited to situations when L1 and L2 items were closely matched in terms of their fine-grained acoustic–phonetic properties (Ju and Luce, 2004) or found no evidence of L2 activation in L1 at all (Weber and Cutler, 2004). This apparent asymmetry has often been attributed to lowered language proficiency in L2 compared to L1. Lowered proficiency would presumably lead to a reduction of the subjective frequency of L2 items, leaving them less likely to compete with L1 items in language processing. A further consequence of low L2 proficiency emerges from the Revised Hierarchical Model (RHM; Kroll and Stewart, 1994). The RHM assumes that lexico-conceptual links follow a more direct route for L1 items, whereas L2 items (especially for low-proficient bilinguals) more often need to rely on L2-to-L1 lexical links in order to access conceptual features. Hence, the L1 translation equivalent is often explicitly accessed in L2 word recognition, whereas L1 processing need not involve L2 lexical activation.

In summary, it is clear that there is a great deal of evidence for between language lexical activation in L2 processing, both in the visual and in the auditory domains. However, relative language proficiency as well as the degree of fine-grained acoustic similarity between languages may serve to temper the activation of L2 lexical candidates in L1 processing. Interestingly, the degree of cross-linguistic activation in L2 listening may, itself, not be immune to modulating influences as has recently been demonstrated by studies that investigated cross-linguistic lexical activation using full sentence contexts.

1.3. *Influence of sentence context*

The majority of studies that demonstrated cross-linguistic lexical activation in bilingual speech comprehension used target words embedded in invariant contexts (e.g., "Click on the desk"). However, those few studies that investigated bilingual speech comprehension in semantically rich sentences seem to suggest that bilinguals may be able to utilise contextual constraints to modulate the availability of between-language lexical competitors. Li and Yip (1998), for example, investigated the processing of interlingual homophones in Chinese–English bilinguals. The homophones were pronounced according to Chinese phonetics and embedded in semantically biasing or neutral Chinese sentences. The contextual bias was towards the non-context (English) language meaning of the homophone. The bilinguals' task was to identify visual probes that could be either the English version of the homophone, a phonologically related Chinese word, or a non-homophone control in either English or Chinese. Identification latencies for the English version of the homophone were significantly faster than non-homophone controls in semantically biasing sentences, but not in neutral sentences. This seems to indicate that the bilinguals were able to utilise contextual constraints to enhance the availability of non-target language lexical candidates. However, there have also been reports of reduced between-language lexical activation in studies of auditory sentence comprehension. FitzPatrick and Indefrey (2010) found evidence of within-, but not between language lexical competition in bilingual spoken sentence comprehension. The study focused on the N400, an extensively researched ERP component. The amplitude of the N400 is widely held to index the ease of semantic integration (Brown and Hagoort, 1993; Kutas and Hillyard, 1984), and the peak and onset latency of the N400 component have been shown to be sensitive to the point at which a semantic incongruity is detected (O'Rourke and Holcomb, 2002; Praamstra et al., 1994). FitzPatrick and Indefrey (2010) presented Dutch–English bilinguals with English (i.e., second language) sentences that terminated in a word that was (a) semantically fitting, (b) semantically incongruent, (c) semantically incongruent, but initially congruent due to sharing initial phonemes with the most probable sentence completion within the L2, or (d) the L1 translation equivalent of the most probable sentence completion. The authors reasoned that, if the target word would be initially perceived as congruent, the N400 would be delayed in comparison to non-overlapping semantically incongruent target words. Whereas intralingual (i.e., within L2) phonemic overlap led to a significantly delayed N400, interlingual (i.e., L1–L2) phonemic overlap did not, suggesting that within-, but not between-language lexical competitors are available for semantic integration. Similarly, in a visual word paradigm using interlingual homophones, Chambers and Cooke (2009) observed reduced between language lexical competition, when sentence contexts were constrained towards the target language interpretation of the homophone, compared to sentences where the context provided no disambiguating information. These findings converge with observations in the domain of bilingual visual word recognition which show a marked reduction of between language lexical activation in sentence reading (Duyck et al., 2007; Libben and Titone, 2009; Schwartz and Arêas Da Luz Fontes, 2008; Schwartz and Kroll, 2006; Van Hell and De Groot, 2008). Duyck et al. (2007) presented bilinguals with between language cognates, near-cognates, and non-cognate control words in isolation and in sentences presented word-by-word. Recognition of both cognates and near-cognates was facilitated compared to control words; however the near-cognate effect disappeared when the full sentence was presented while the cognate effect remained. The authors concluded that the presence of a sentence context "may influence, but does not nullify" cross-linguistic lexical activation. This assertion is supported by observations by Schwartz and Arêas Da Luz Fontes (2008) who investigated the effect of context on between-language mediated form priming and -semantic priming. Using Spanish–English bilinguals, in an English (L2) only task, they obtained priming when the prime (*bark*) had a form mediated relationship via the L1 (*barco*; Spanish for 'boat') to the target (BOAT) for isolated words, but not when the words were embedded in sentence contexts (e.g., "The baby woke up every time the dog would *bark*"). Further evidence suggests an effect of semantic constraint on cross-linguistic lexical activation. Schwartz and Kroll (2006) (see also Van Hell and De Groot (2008)) observed a cognate facilitation effect for the reading of cognates in low-constraint sentences, but this effect was substantially reduced when the cognates were presented in high-constraint sentence contexts.

There remains some controversy surrounding the nature of the interaction between the bilingual lexical activation process and top-down influences of the sentence context, as well as the time course at which this interaction might manifest itself. On the one hand, some studies find no evidence of between-language activation in semantically constraining sentence contexts (Schwartz and Kroll, 2006; Van Hell and De Groot, 2008), although this might only be the case for non-identical cognates (Duyck et al., 2007). On the other hand, there are suggestions that contextual effects might only appear late in the word recognition process after initial language non-selective lexical access. Libben and Titone (2009), for instance, investigated the influence of high, and low sentential semantic constraint on the reading of interlingual homographs, cognates, and matched control words using an eye-tracking paradigm. Particularly the First Fixation Duration (FFD) and First Pass Gaze Duration (GD), measures that are assumed to reflect the initial stages of lexical activation, see Rayner (1998), showed facilitation for cognates and interference effects for interlingual homographs, relative to control words, independent of sentential constraint. Interestingly, measures of later processing stages such Go-past Time (GPT) and Total Reading Time (TRT) showed interference effects for homophones and facilitation for cognates, but only in low constraint sentences. Still other observations seem to show that contextual constraints might also modulate between-language lexical activation in L1 reading. Titone et al. (2011), for instance, showed that cognate facilitation on the GPT and TRT measures was reduced in high constraint sentences, but only when the experiment did not contain fillers from the non-target language. These results might be taken to suggest that contextual constraints assert themselves after initial language non-selective lexical access. However, results obtained by Van Assche et al. (2010), who observed cognate facilitation on the FFD, GD, and GPT measures in both high and low-constraint sentences, cast doubt on that assertion.

Taken together, findings in both bilingual speech comprehension as well as bilingual visual word recognition suggest that the sentence context might modulate the availability of between-language lexical candidates during bilingual language processing. However questions remain surrounding, the extent to which between language lexical activation can take place even in semantically constraining context, as well as the time-course with which contextual constraints can exert themselves.

1.4. Objectives

The present study aims to shed more light on the nature of sentence context effects on between-language lexical activation in both the native, and the non-native language of highly proficient, late onset bilinguals. In particular, we aim to uncover how sentential bias affects the time course of between-language ambiguity resolution. As in our previous study (FitzPatrick and Indefrey, 2010) we will exploit the sensitivity of the N400 to the moment at which a semantic incongruity becomes apparent to the listener. Table 1 shows examples of the stimulus materials used in all 4 experiments that are described below.

In Experiment 1 of the present study, Dutch native speakers listened to spoken sentences in their second language (English). The sentences included an interlingual near-homophone (e.g., *pet*, meaning 'hat' in Dutch). In different conditions the sentence context was (1) semantically biased towards the target language (English; L2) meaning of the homophone (*target biased*), while its non-target language (Dutch; L1) meaning did not semantically fit the sentence context, (2) semantically biased towards the non-target language (Dutch; L1) meaning of the homophone (*non-target biased*), (3) semantically incongruent with respect to both the target and the non-target language meaning of the homophone (*fully incongruent*), or (4) semantically fitting with a non-homophone (e.g., *kite*, Dutch: 'vlieger') control word (*fully congruent*). We reasoned that the activation of the non-target language semantics of homophones would lead to a decreased N400 (relative to *fully incongruent* words) in the *non-target biased* condition due to the presence of semantically fitting semantic features (from the non-target language), and an increased N400 (relative to *fully congruent* words) in the *target biased* condition due to the presence of semantically incongruent features (from the non-target language). By contrast, absence of between-language meaning activation would mean that the *target biased* condition should not show an N400 effect (relative to the *fully congruent* condition), and the *non-target biased* condition should show the same N400 magnitude as the *fully incongruent* condition.

The potential activation of cross-linguistic semantics in Experiment 1 raises a caveat for our interpretation, as it has

Table 1 – Examples of stimulus materials.

Condition	*Lead-in sentence*	*Target*
Experiment 1 (target language L2):		
Target biased	My cat is my favourite	*pet* (NL: 'hat')
Non-target biased	The policeman wore a	*pet* (NL: 'hat')
Fully incongruent	Jeremy drove the	*pet* (NL: 'hat')
Fully congruent	We went to the Vatican to see the	*Pope* (NL: 'paus')
Experiment 2 (target language L2):		
Fully congruent	The rose grew on a thorny	*bush* (NL: 'struik')
Fully incongruent	The broken glass gave him a nasty	*salt* (NL: 'zout')
Fitting switch	For balance, the cat has a	*staart* (EN: 'tail')
Non-fitting switch	You wear your watch on your	*draad* (EN: 'thread')
Experiment 3 (target language L1):		
Target biased	Als je verhuist, stop je al je boeken in een	*doos* (EN homophone: 'dose')
	When you move house, you put all your books in a	*box*
Non-target biased	De arts had het juiste medicijn, maar de onjuiste	*doos* (EN homophone: 'dose')
	The doctor had the correct medicine, but the incorrect	*box*
Fully incongruent	De voetballer schoot de	*doos* (EN homophone: 'dose')
	The footballer shot the	*box*
Fully congruent	De schoolkinderen mochten buiten spelen op het	*plein*
	The school children were allowed to play outdoors on the	*square*
Experiment 4 (target language L1):		
Fully congruent	Het kind zat bij moeder op	*schoot*
	The child was sitting on mother's	*lap*
Fully incongruent	Voor de wedstrijd, gooit de scheidsrechter een	*kuil*
	Before the match, the referee tosses a	*hole*
Fitting switch	Rotte eieren hebben een hele vieze	*smell* (NL: 'geur')
	Rotten eggs have a very foul	*smell*
Non-fitting switch	Het meest opvallende kenmerk van een olifant is zijn	*lawn* (NL: 'gazon')
	The most salient characteristic of an elephant is its	*lawn*

been shown that language switches can have a substantial effect on the electrophysiological waveform (Martin et al., 2009; Moreno et al., 2002; Van Der Meij et al., 2010; Van Hell and Witteman, 2009). Moreno et al. (2002) observed that semantically congruent language switches (e.g., "He heard a knock at the *puerta*") did not result in an N400 effect compared to semantically congruent control sentences (e.g., "He heard a knock at the *door*") whereas lexical switches (e.g., "He heard a knock at the *entrance*") did. Instead the language switches elicited a Late Positive Component (LPC), which the authors speculate may be a reflection of the perception of a language switch. In our Experiment 1, a modulation of the N400 could potentially be attributable to an LPC effect that temporally overlaps with the N400. We therefore considered it important to compare results from Experiment 1 to a situation in which the activation of cross-linguistic semantics was combined with an explicit language switch. Thus we conducted a second experiment, in which the same participants were presented with sentences containing semantically congruent and incongruent language switched critical words instead of homophones. We expected language switches (irrespective of semantic congruity) to elicit LPC effects compared to semantically congruent and incongruent non-switches. By comparing the amplitudes, latencies, and scalp topographies of these LPC effects to modulations of the N400 obtained in Experiment 1 we can, thus, establish whether these modulations reflect true modifications of the N400, or temporally co-occurring N400 and LPC effects.

While cross-linguistic semantic activation is frequently found in non-native language processing, findings of L2 activation during L1 processing have been more inconsistent (for an overview, see 1.2). In Experiments 3 and 4 we therefore set out to investigate whether the activation of non-target language semantics would lead to N400 modulations for *target biased*, and *non-target biased* homophones (Experiment 3) and/or LPC effects for semantically *congruent*, or *incongruent, language switches* (Experiment 4). Experiments 3 and 4 tested an independent group of Dutch–English bilinguals and employed the stimulus materials from Experiments 1 and 2, which were translated into Dutch (L1). If indeed non-target language activation is less extensive in native-language processing we would predict to find smaller modulations of the N400 (versus *fully congruent* controls) for *target biased*, and *non-target biased* homophones, compared to Experiment 1, and smaller LPC effects for *language switches* compared to Experiment 2.

2. Results

2.1. Experiment 1

The waveforms for *semantically incongruent* and *non-target bias* conditions show an increased negativity in the 300–800 ms latency range relative to the *fully congruent* condition (Figs. 1, 2 and S1). This negativity has an early offset in the *non-target bias* condition (compared to the *fully incongruent* condition).

2.1.1. Time window analyses

In the 250–450 ms time window (Figs. 3 and S2) the ANOVA yielded a significant main effect of condition ($F(3,72)=5.704$, $p<0.01$, $\eta^2=0.192$). A priori contrasts revealed significant differences between the *fully congruent* and the *fully incongruent* ($F(1,24)=6.016$, $p<0.05$, $\eta^2=0.200$), and *non-target bias* ($F(1,24)=9.945$, $p<0.01$, $\eta^2=0.293$) conditions, with the latter conditions both exhibiting a greater negativity than the former. There was also a significant condition by site interaction ($F(24,576)=3.452$, $p<0.05$, $\eta^2=0.126$) reflecting, firstly, the fact that the difference between the *fully congruent* and *fully incongruent* conditions reached significance on all sites except the right-occipital site, and secondly, the fact that that the difference between the *fully congruent* and *non-target bias* conditions reached significance on all sites except the left-precentral, left-occipital, and right-occipital sites.

In the 450–650 ms time window (Figs. 3 and S3) the ANOVA yielded a significant main effect of condition ($F(3,72)=9.546$, $p<0.001$, $\eta^2=0.285$), and a significant main effect of site ($F(8,192)=6.007$, $p<0.01$, $\eta^2=0.200$). A priori contrasts revealed significant differences between the *fully congruent* and the *fully incongruent* ($F(1,24)=16.689$, $p<0.001$, $\eta^2=0.410$) conditions, with the *fully incongruent* condition exhibiting a greater negativity than the *fully congruent* condition. None of the other contrasts yielded significant results (all $F<2.1$, $p=n.s.$).

In the 650–850 ms time window (Figs. 3 and S4) the ANOVA yielded a non-significant trend towards a main effect of condition ($F(3,72)=2.502$, $p<0.08$, $\eta^2=0.094$) and a significant main effect of site ($F(8,192)=8.066$, $p<0.01$, $\eta^2=0.252$). A priori contrasts revealed significant differences between the *fully congruent* and the *target bias* ($F(1,24)=5.791$, $p<0.05$, $\eta^2=0.194$) conditions. With the *target bias* condition exhibiting a greater negativity than the *fully congruent* condition. There was also a non-significant trend towards a difference between the *fully incongruent* and *fully congruent* conditions ($F(1,24)=3.619$, $p<0.07$, $\eta^2=0.131$) conditions. None of the other contrasts yielded significant results (all $F<1$, $p=n.s.$).

2.1.2. Cluster randomisation statistics

Fig. 4 shows the topographical distribution of the significant clusters. The contrast between the *fully incongruent condition* and the *fully congruent* condition yielded a significant negative effect ($p<0.001$) emerging at 424 ms after critical word onset and lasting until 716 ms. Contrasting *non-target bias* words with *fully congruent* words yielded a significant negative effect ($p<0.05$) starting at 374 ms after critical word onset. However, different from homophones in the *fully incongruent* condition, the negativity to homophones in the *non-target bias* condition offset nearly 300 ms earlier at 440 ms. Contrasting homophones in the *target bias* condition with the *fully congruent* condition yielded a significant negativity ($p<0.05$) at 742 ms lasting until 834 ms.

2.2. Experiment 2

The waveforms for *fully incongruent*, *fitting-switch* and *non-fitting switch* conditions show an increased negativity in the 300–800 ms latency range relative to the *fully congruent* condition (Figs. 1, 2 and S5). This negativity has an early offset in the fitting-switch condition (compared to the *fully incongruent*

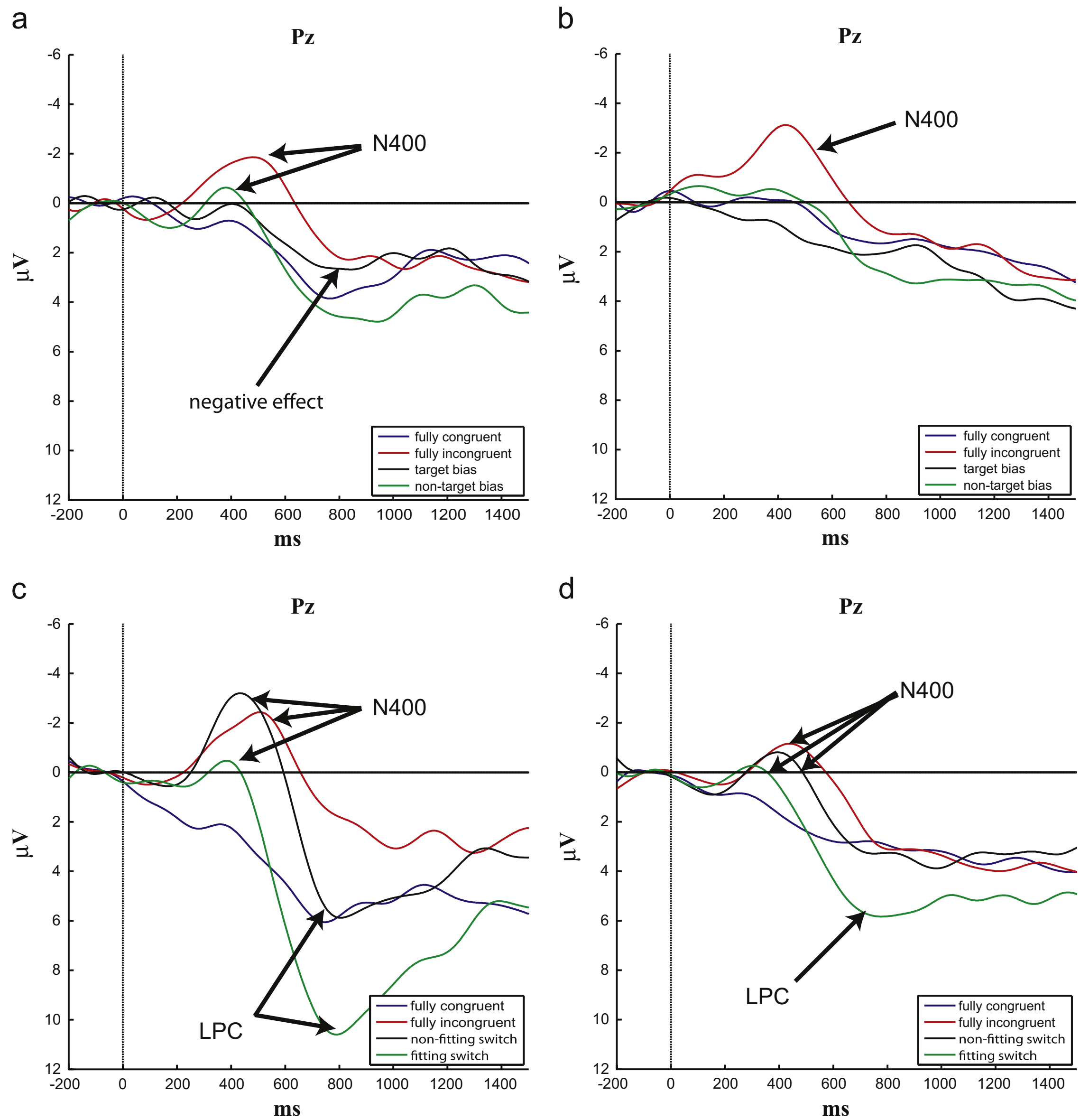

Fig. 1 – ERP waveforms on the Pz electrode for (a) homophones in L2 sentence contexts, (b) homophones in L1 sentence contexts, (c) language switches in L2 sentence contexts, and (d) language switches in L1 sentence contexts. Waveforms were filtered with a 5 Hz low-pass filter for presentation purposes only.

and *non-fitting switch* conditions) and is most pronounced on centro-parietal electrodes. Both the fitting switch and non-fitting switch show a late positive deflection in the ERP waveforms at around 700 ms.

2.2.1. *Time window analyses*

In the 250–450 ms time window (Figs. 3 and S2) the ANOVA yielded a significant main effect of switching ($F(1,24)=14.058$, $p<0.01$, $\eta^2=0.369$). However, a significant interaction between congruity and switching ($F(1,24)=5.437$, $p<0.05$, $\eta^2=0.185$) revealed that this was due to a difference between semantically congruent and incongruent words which only emerged when they were not language switched. There was also a significant interaction between switching and site ($F(8,192)=7.738$, $p<0.01$, $\eta^2=0.244$), reflecting the fact that most sites showed an increased negativity to switched versus non-switched words with the exception of the left occipital site.

In the 450–650 ms time window (Figs. 3 and S3) the ANOVA yielded a significant main effect of congruity ($F(1,24)=46.727$, $p<0.001$, $\eta^2=0.661$), with semantically incongruent words showing a greater negativity than semantically congruent words.

In the 650–850 ms time window (Figs. 3 and S4) the ANOVA yielded a significant main effect of congruity ($F(1,24)=24.182$, $p<0.001$, $\eta^2=0.502$), with semantically incongruent words showing a greater negativity than semantically congruent words. There was also a significant main effect of switching ($F(1,24)=18.510$, $p<0.001$, $\eta^2=0.435$), with language switched words showing a greater positivity than non-switched words.

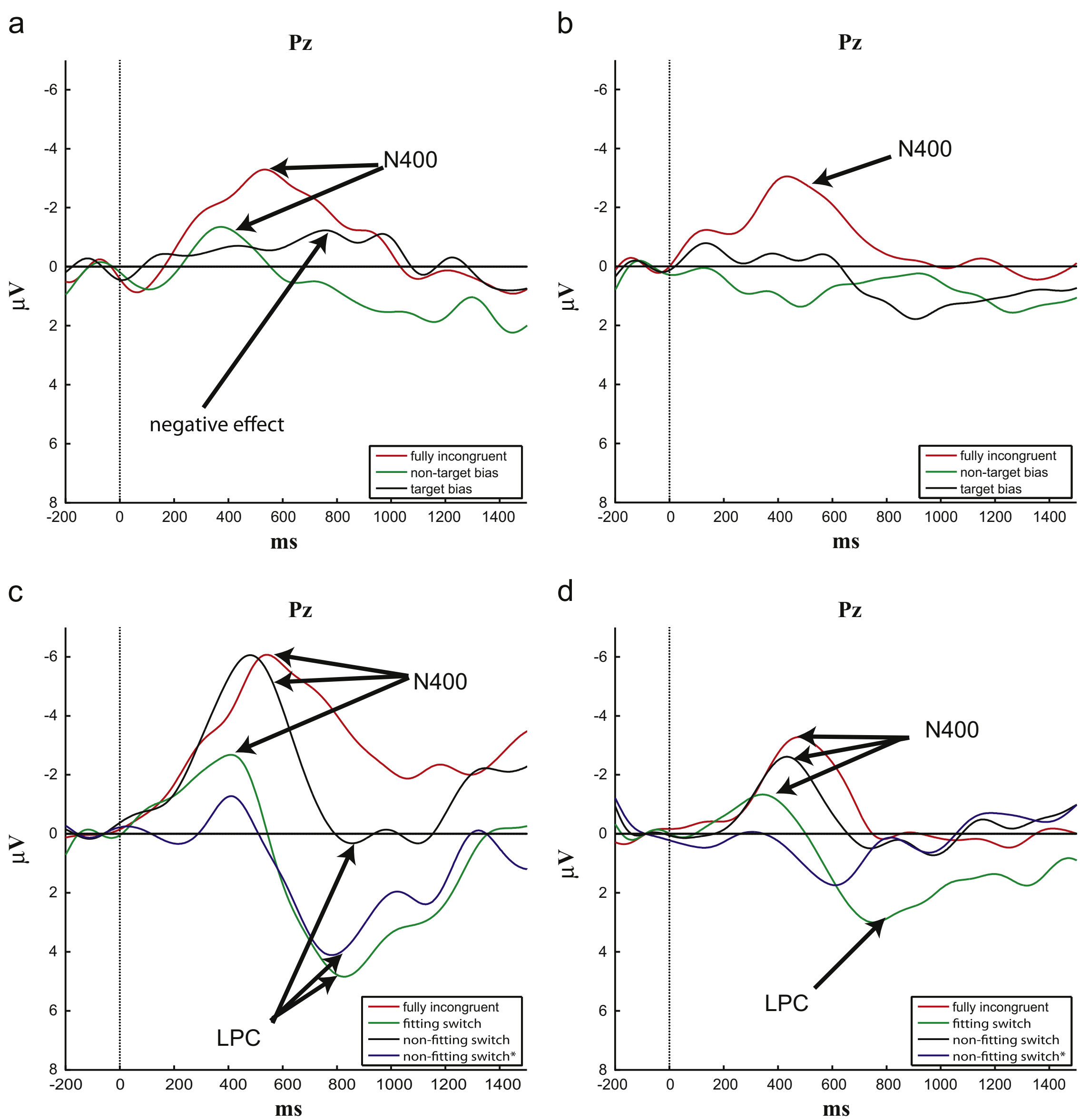

Fig. 2 – Difference waveforms (fully congruent condition subtracted) on the Pz electrode for (a) homophones in L2 sentence contexts, (b) homophones in L1 sentence contexts, (c) language switches in L2 sentence contexts, and (d) language switches in L1 sentence contexts. Waveforms were filtered with a 5 Hz low-pass filter for presentation purposes only. Asterisks denote subtractions of fully incongruent condition.

2.2.2. *Cluster randomisation statistics*

Fig. 4 shows the topographical distribution of the significant clusters. Semantic incongruity led to a negative effect irrespective of whether the critical word was a language switch ($p<0.001$; onset 296 ms; offset 754 ms) or not ($p<0.001$; onset 278 ms; offset 876 ms). Language switched, semantically congruent words led to a transient negativity ($p<0.001$; onset 202 ms; offset 560 ms). Language switches additionally elicited a late positive effect irrespective of whether they were semantically congruent ($p<0.001$; onset 572 ms; offset 1112 ms) or semantically incongruent ($p<0.05$; onset 598 ms; offset 1178 ms). The isovoltage topographical plots (Fig. 2) of the negative effect show it to have a centro-parietal distribution, whereas the positive effect had a largely parieto-occipital distribution. The non-switched, semantically incongruent words did not exhibit the positive effect.

2.3. *Experiment 3*

The waveform for *semantically incongruent* condition shows an increased negativity in the 300–800 ms latency range relative to the *fully congruent* condition (Figs. 1, 2 and S6).

2.3.1. *Time window analyses*

In the 250–450 ms time window (Figs. 3 and S2) the ANOVA yielded a significant main effect of condition ($F(3,96)=6.682$, $p<0.01$, $\eta^2=0.173$). A priori contrasts revealed significant differences between the *fully congruent* and the *fully incongruent* ($F(1,32)=6.408$, $p<0.05$, $\eta^2=0.167$) condition. There was also a non-significant trend towards a difference between the *fully congruent* condition and the *target bias* condition ($F(1,32)=4.018$, $p<0.06$, $\eta^2=0.112$), with the *target bias* condition exhibiting a more positive going waveform than the *fully congruent*

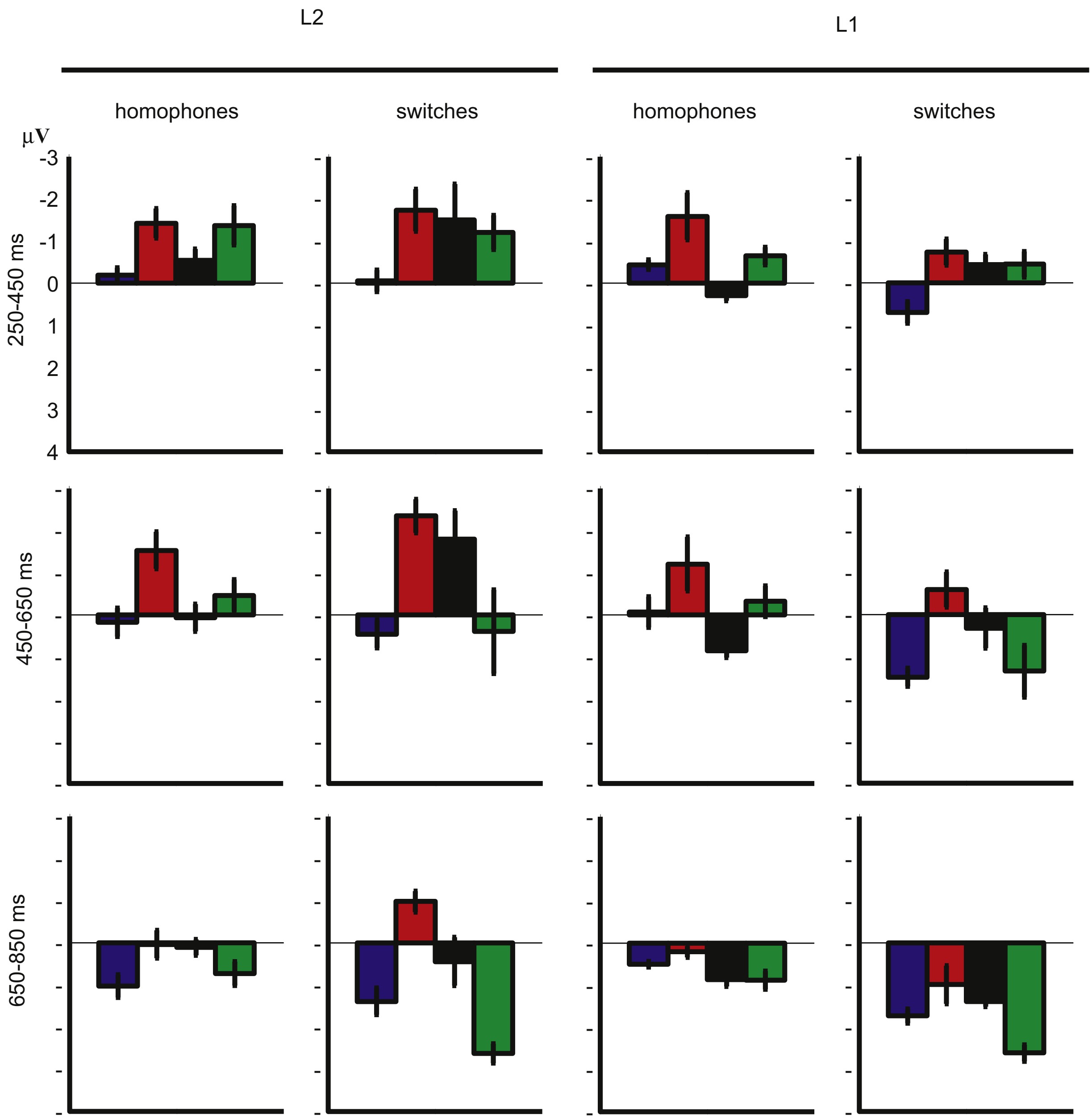

Fig. 3 – Mean amplitudes of the global field potentials in the 250–450 ms (top), 450–650 ms (middle), and 650–850 ms (bottom) time windows. Left and centre left plots (language = L2) correspond to Experiments 1 (homophones) and 2 (switches); Right and centre right plots (language = L1) correspond to Experiments 3 (homophones) and 4 (switches). Error bars denote standard deviations.

condition. There was also a significant condition by site interaction ($F(24,768)=3.726$, $p<0.01$, $\eta^2=0.104$) reflecting the fact that the difference between the *fully congruent* and *fully incongruent* conditions failed to reach significance on left- and right occipital, as well as, left precentral sites (all $F(1,32)<2$).

In the 450–650 ms time window (Figs. 3 and S3) the ANOVA yielded a significant main effect of condition ($F(3,96)=6.269$, $p<0.01$, $\eta^2=0.164$). A priori contrasts revealed significant differences between the *fully congruent* and the *fully incongruent* ($F(1,32)=6.143$, $p<0.05$, $\eta^2=0.161$) condition. There was also a non-significant trend towards a difference between the *fully congruent* condition and the *target bias* condition ($F(1,32)=3.593$, $p<0.07$, $\eta^2=0.101$), with the target bias condition exhibiting a more positive going waveform than the *fully congruent* condition. There was also a significant condition by site interaction ($F(24,768)=4.656$, $p<0.001$, $\eta^2=0.127$) reflecting the fact that the difference between the *fully congruent* and *fully incongruent* conditions failed to reach significance on right-frontal, left-precentral, left-postcentral, and right-occipital sites (all $F(1,32)<2$).

In the 650–850 ms time window (Figs. 3 and S4) the ANOVA yielded a significant main effect of site ($F(8,256)=7.827$, $p<0.01$, $\eta^2=0.197$) and a significant condition by site interaction ($F(24,768)=4.192$, $p<0.01$, $\eta^2=0.116$) reflecting the fact

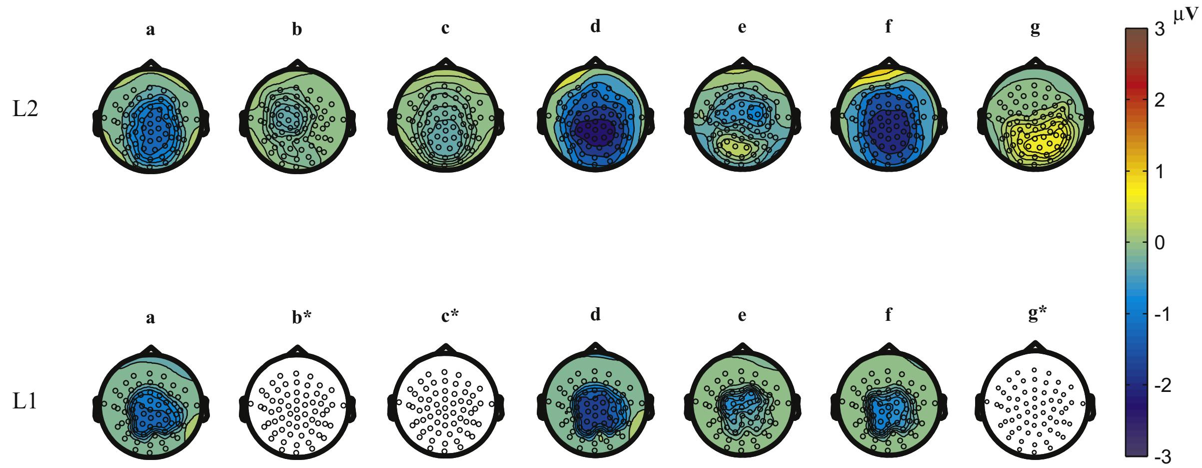

Fig. 4 – Topographical plots for the significant clusters in Experiments 1 and 2 (L2; top row) and Experiments 3 and 4 (L1; bottom row) between the (a) *fully congruent* & *incongruent homophones*, (b) *fully congruent* & *target bias*, (c) *fully congruent* & *non-target bias*, (d) *fully congruent* & *fully incongruent*, (e) *fully congruent* & *congruent, language switch*, (f) *fully congruent* & *incongruent, language switch*, and (g) *fully incongruent* & *incongruent, language switch* conditions. Asterisks denote contrasts where no significant clusters emerged.

that right-postcentral and left-occipital sites continued to show a difference between the *fully incongruent* and *fully congruent* conditions, while other sites did not.

2.3.2. *Cluster randomisation statistics*

Fig. 4 shows the topographical distribution of the significant clusters. The contrast between the *fully incongruent* condition and the *fully congruent* condition yielded a significant negative effect ($p<0.01$) emerging at 314 ms after critical word onset and lasting until 622 ms. Contrasts between the *target bias* and *non-target bias* conditions and the *fully congruent* condition did not yield any significant effects.

2.4. *Experiment 4*

The waveforms for *fully incongruent*, *fitting-switch* and *non-fitting switch* conditions show an increased negativity in the 300–800 ms latency range relative to the *fully congruent* condition (Figs. 1, 2 and S7). This negativity has an early offset in the fitting-switch condition (compared to the *fully incongruent* and *non-fitting switch* conditions) and is most pronounced on centro-parietal electrodes. Both the fitting switch and non-fitting switch show a late positive deflection in the ERP waveforms at around 700 ms.

2.4.1. *Time window analyses*

In the 250–450 ms time window (Figs. 3 and S2) the ANOVA yielded a significant main effect of congruity ($F(1,32)=6.268$, $p<0.05$, $\eta^2=0.163$), with semantically incongruent conditions exhibiting more negative ERP waveforms than semantically congruent conditions. However, a significant switching by congruity interaction ($F(1,32)=9.587$, $p<0.01$, $\eta^2=0.231$) revealed that the congruity effect was only present for non-switched words. There was also a significant main effect of site ($F(8,256)=9.597$, $p<0.001$, $\eta^2=0.231$) reflecting a greater negativity on centro-posterior electro-sites than on other sites.

In the 450–650 ms time window (Figs. 3 and S3) the ANOVA yielded a significant main effect of congruity ($F(1,32)=17.473$, $p<0.001$, $\eta^2=0.353$), with semantically incongruent conditions exhibiting more negative ERP waveforms than semantically congruent conditions. There was also a significant main effect of site ($F(8,256)=16.965$, $p<0.001$, $\eta^2=0.346$) reflecting a greater negativity on centro-posterior electro sites than on other sites. There was a significant switching by site interaction ($F(8,256)=5.605$, $p<0.01$, $\eta^2=0.149$), with posterior electrode sites exhibiting a larger positivity for language-switched words than for non-switched words. There was also a significant congruity by site interaction ($F(8,256)=3.676$, $p<0.05$, $\eta^2=0.103$) with mid-line electrodes exhibiting a larger congruity effect than other sites. Finally, there was a significant three-way interaction between congruity, site, and switching ($F(8,256)=4.823$, $p<0.01$, $\eta^2=0.131$) indicating that the congruity by switching interaction was different across electrode sites.

In the 650–850 ms time window (Figs. 3 and S4) the ANOVA yielded a significant main effect of congruity ($F(1,32)=6.031$, $p<0.05$, $\eta^2=0.159$), with semantically incongruent conditions exhibiting more negative ERP waveforms than semantically congruent conditions. There was also a significant main effect of site ($F(8,256)=13.654$, $p<0.001$, $\eta^2=0.299$) reflecting a greater negativity on centro-posterior electro-sites than on other sites. There was a significant switching by site interaction ($F(8,256)=6.183$, $p<0.01$, $\eta^2=0.162$), with posterior electrode sites exhibiting a larger positivity for language-switched words than for non-switched words. Finally, there was a significant three-way interaction between congruity, site, and switching ($F(8,256)=11.501$, $p<0.001$, $\eta^2=0.264$) indicating that the congruity by switching interaction was different across electrode sites.

2.4.2. *Cluster randomisation statistics*

Fig. 4 shows the topographical distribution of the significant clusters. Semantic incongruity led to a negative effect irrespective of whether the critical word was a language switch ($p<0.01$; onset 334 ms; offset 532 ms) or not ($p<0.001$; onset 328 ms; offset 694 ms). Language switched, semantically congruent words led to a transient negativity ($p<0.001$; onset 206 ms; offset 488 ms) and a late positive effect ($p<0.05$; onset 722 ms; offset 816 ms).

3. Discussion

3.1. Experiment 1

Experiment 1 investigated the time course of cross-linguistic semantic activation in L2 spoken sentence contexts by manipulating the semantic fit of interlingual homophones. As expected, results revealed an N400 effect between homophones in the *fully incongruent* condition and *fully congruent* words. Both the onset and the offset of the N400 effect were later than what is typically observed in studies of first language comprehension that use N400 paradigms (cf., Hagoort and Brown, 2000). However, similar N400 delays have been reported for L2 versus L1 speech comprehension (e.g., Ardal et al., 1990; Hahne, 2001). Significant negative effects also emerged between firstly, the *non-target bias* condition and the *fully congruent* condition, and secondly the *target bias* condition and the *fully congruent* condition. Both these effects exhibited different time-courses from the N400 effect to *fully incongruent* words. Specifically the *non-target bias* words elicited an N400 effect in an early time window, which was absent in later time windows. Conversely, the *target bias* condition elicited a negative effect[1] in a late time window, which was absent in earlier time windows. The results from Experiment 1 seem to show cross-linguistic semantic activation of homophone meanings. If only the semantic features of the target language meaning of the homophone would have been considered, the *non-target bias* condition would have elicited a canonical N400 effect, and the *target bias* condition would not have elicited a negative effect. The time course of the negative effect suggests a target language priority for homophone meaning activation. The semantic features of the non-target language seem not to be initially available (cf., FitzPatrick and Indefrey, 2010) leading to an initial incongruity in the *non-target bias* condition, and initial congruity in the *target bias* condition. But when the non-target language semantics assert themselves after a short delay this then leads to an attenuation of the negative effect in the *non-target bias* condition (compared to the *fully incongruent* condition) due to the congruity of the non-target language semantics, and an emergent negative effect in the *target bias* condition (relative to the *fully congruent* condition) due to the incongruity of the non-target language semantics. In order to properly interpret these results, it is important to establish whether these data reflect a modulation of the N400 component or could be attributed to superposition of an N400 and an LPC (e.g., Moreno et al., 2002). Thus, our aim for experiment 2 was to compare results from experiment 1 to a situation in which activation of cross-linguistic semantics to target words is a given, namely when using language switches.

3.2. Experiment 2

In Experiment 2 we investigated the electrophysiological correlates of perceived language switching in spoken language comprehension using semantically congruent and incongruent language switches. Processing of semantically incongruent words led to an N400 effect compared to semantically congruent words, irrespective of whether the words were language switched or not. When the language switch was semantically congruent with the sentence context we observed a transient N400 effect compared to *fully congruent* words. This effect was both smaller and of shorter duration than the N400 effects observed for semantically incongruent (switched and non-switched) words. This observation differed from Moreno et al. (2002) who did not find an N400 effect for language-switched words relative to non-switched words (but see Van Der Meij et al. (2010)), but rather observed that language switches elicited a left-frontal negativity. Moreno et al. (2002) speculate that the early effect might reflect a Left Anterior Negativity (LAN) which arises due to working memory demands resulting from syntactically integrating the language switched word into the base language context. In our data the scalp distribution of the transient negative effect in the contrast between the *fitting switch* condition and the *fully congruent* condition was identical to that of the N400, which leaves us with little indication that this negative effect could be distinct from the N400. The fact that the contrast between the *fitting switch* condition and the *fully congruent* condition yielded an N400 effect at all, suggests that the (semantically fitting) L1 words in the *fitting switch* condition are initially unavailable for semantic integration, but are nevertheless accessed after a short delay. This observation bears a strong resemblance to the pattern of results in Experiment 1, which showed a transient N400 in the *non-target bias* condition and a later negative effect in the *target bias* condition.

Language switches also elicited an LPC with a mostly posterior scalp distribution. The scalp topography and latency of the LPC effect (Figs. 1, 2, 4, S4 and S5) seem to be distinct from those of the N400 modulation we observed in Experiment 1. In light of these results we now return to the question of whether the N400 modulation, which was observed in Experiment 1, could be attributable to an LPC effect, due to cross-language semantic activation to homophones. A number of different observations render this interpretation unlikely. Firstly, the magnitude of the LPC to language switches in Experiment 2 (Figs. 1 and 2) was far greater (~6 μV) compared to the positive deflection of the waveform in the non-target bias condition of Experiment 1 (~1.5 μV). Secondly, the LPC showed a largely posterior scalp

[1]We purposefully refrain from referring to the latter negativity as an N400 as, whereas the scalp topography of the N400 effect in the *non-target bias* condition was almost identical to that in the *fully incongruent* condition, the negativity in the *target bias* condition had a more left-frontal distribution (Figs. 4 and S1). These differences could either signify a separate ERP component, part of a multiphasic N400 component, or an N400 effect that has been modulated by processes that would otherwise have been obscured by the main N400 effect. However, the present data cannot distinguish between these accounts.

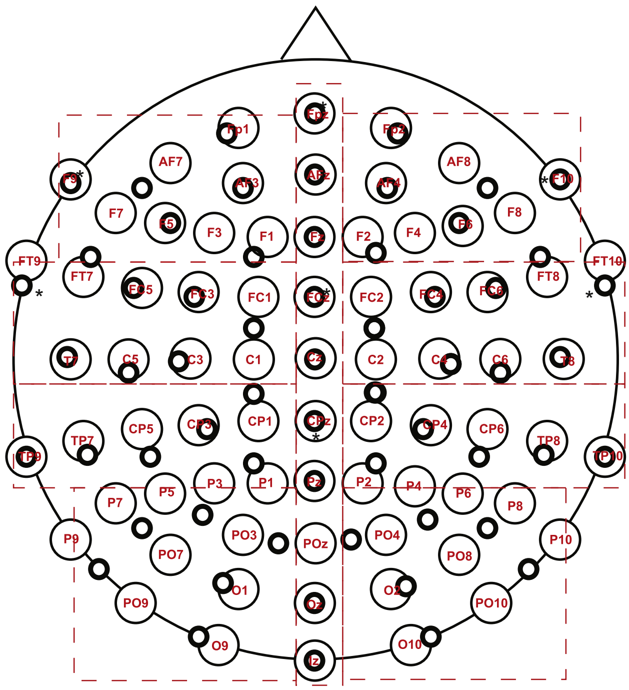

Fig. 5 – Radial projection of electrode positions of the equidistant placement system (small black circles) relative to the 10–20 system. Boxes denote regions for statistical analyses. Asterisks denote electrodes excluded from analysis.

topography (Figs. 4 and S4) whereas the topography of the negativity in the non-target bias condition was broadly similar to that of the N400 effect in the *target bias* condition (Figs. 4 and S2). Finally, the *non-fitting switch* condition in Experiment 2 was indicative of a superposition of an LPC onto an on-going N400 effect. In this condition the positive deflection starts to emerge around 600 ms post-stimulus, whereas the negativity in the *non-target bias* condition of Experiment 1 had already offset around 440 ms post-stimulus. In summary, the results from Experiment 2 show a largely posterior LPC effect to language switches irrespective of their semantic congruity. The scalp distribution and latency of this effect seem to be distinct from the observed modulation of the N400 in Experiment 1. We therefore consider it unlikely that the N400 modulation in Experiment 1 is the result of co-occurring N400 and LPC effects.

3.3. *Experiment 3*

Experiment 3 investigated the time course of cross-linguistic semantic activation in native-language, spoken sentence contexts by manipulating the semantic fit of interlingual homophones. Surprisingly (and in contrast to Experiment 1), only the *fully incongruent* condition exhibited a significant N400 effect compared to the *fully congruent* condition. This would seem to suggest that the homophones were treated as semantically congruent words in both the *target bias* and the *non-target bias* conditions, which in turn means that both homophone meanings were immediately available to the bilingual. We also failed to observe a significant LPC effect in any of our homophone conditions in Experiment 3. This observation strengthens our assertion that the observed time-course differences in Experiment 1 were not due to the superposition of an N400 effect and an LPC effect.

The finding that both homophone meanings are immediately available to the bilingual in L1 listening would seem at odds with the target-language priority observed in Experiments 1 and 2. One possible account for this pattern of results entails an asymmetry of non-target language availability in L1 processing on the one hand, and L2 processing on the other. Such an account would be compatible with observations that, for unbalanced bilinguals, switching into a dominant language incurs greater processing costs than switching into a non-dominant language (e.g., Costa and Santesteban, 2004; Meuter and Allport, 1999). This asymmetry is usually explained by the existence of a greater need to inhibit the L1 (or dominant language) during L2 processing than there is to inhibit the L2 during L1 processing. If a similar asymmetry of non-target language suppression exists in the spoken language comprehension domain, one would expect L2 lexical items (whether they be homophones or language switches) to be more readily available in L1 processing than L1 lexical items are in L2 processing. Alternatively, however, the result of Experiment 3 might be due to a special status of non-target language homophones in L1 listening and the processing of language switches might still reveal a target language priority on a par with Experiments 1 and 2.

In Experiment 4 we therefore set out to establish whether also the meanings of language-switched words are immediately available to the bilingual in native-language comprehension. If L2 words are indeed readily available in L1 listening, we expected semantically fitting language switches (*fitting switch*) not to elicit an N400 effect (compared to *fully congruent*, non-switched controls), whereas semantically non-fitting language switches (*non-fitting switch*) should show a similar N400 as semantically incongruent, non-switched words (*fully incongruent* condition). By contrast, if non-homophonic L2 words are only available with a delay we would expect a replication of the results of Experiment 2

3.4. *Experiment 4*

As in Experiment 2, processing of semantically incongruent words led to an N400 effect compared to semantically congruent words, irrespective of whether the words were language switched or not. When the language switch was semantically congruent with the sentence context an initial negativity emerged, followed by an LPC with a mostly posterior scalp distribution. However (in contrast to Experiment 2), the *non-fitting switch* condition failed to show a significant LPC effect compared to the *fully incongruent* condition. A possible reason for this discrepancy is that the LPC effect might be less strong in native-language comprehension than in non-native comprehension. To test this possibility, we computed the area-under-curve (AUC) between the ERP waveforms of the *fully congruent* and *fitting switch* conditions and between the *fully incongruent* and *non-fitting switch* conditions in the 650–850 ms latency range (and over all electrodes) in both L1 (Experiment 4) and L2 (Experiment 2) listening. We subjected the resulting AUC measure to a univariate ANOVA with congruity (2 levels, congruent and incongruent) as within-groups factor and context language (2 levels, L1 and L2) as between-groups factor. We obtained a significant effect of language ($F(1,92)=4.506$, $p<0.05$) showing that the magnitude of the LPC was reduced when L1 was the context language compared to when L2 was the context language.

The functional relevance of the LPC (or P300/ P3 as it is sometimes referred to in studies of language switching) is as yet unclear. There is, however a rich body of literature pertaining to late positive ERP effects (a full discussion of which is beyond the scope of the present article) in both linguistic and non-linguistic tasks and settings (for a review see Polich (2007)). The amplitude of the LPC is purportedly determined by three separate factors (Johnson, 1986): (1) improbability, with more improbable stimuli eliciting larger positivities, (2) meaning, encompassing task complexity and stimulus valence, and (3) transfer, referring to in how much detail the stimulus is evaluated and how much attention is paid to it. Indeed, results from those studies that showed late positivities to language switches (Martin et al., 2009; Moreno et al., 2002) can easily be interpreted within this framework. As Moreno et al. (2002) point out, the language switches in their study could be considered highly improbable events for their participants and thus not very ecologically valid. The improbability thus explains why the LPC emerged to the language switches but not lexical switches in their study. Martin et al. (2009) found that the amplitude of the late positivity was dependent on whether the language switch was task-relevant or not. In their study, Welsh–English bilinguals judged the number of characters of written prime-target pairs in either English or Welsh while ignoring the other language. When attending to English words there was a significant LPC effect between switches and non-switches when the language switches were from Welsh to English, but not when the switches were from English to Welsh. When attending to Welsh, the pattern was reversed. Thus, similar to Moreno et al. (2002) the LPC effect in our study could be explained in terms of the improbable nature of the language switches. In contrast to language switches, the homophones in our Experiments 1 and 3 were members of the task set (i.e., target language), therefore were not considered improbable items, and consequently did not elicit LPC components. On the basis of this interpretation Dutch–English bilinguals may have exhibited a smaller LPC effect when processing Dutch-to-English language switches than when processing English-to-Dutch language switches, due to the fact that it is far more common for the Dutch–English bilinguals to hear English words in Dutch discourse than the reverse.

3.5. *General discussion*

The present study investigated the processing of interlingual homophones and language switches in bilingual spoken language comprehension. In Experiments 1 and 2 the language of the context was English (the participants' L2), and in Experiments 3 and 4 the language of the context was Dutch (the participants' L1). The findings from Experiments 1 and 2 converge on the observation that L1 semantics (both for homophones and for language switches) only become available after a short delay in L2 speech comprehension. Additionally, language switches elicit an LPC component, which is not evident for interlingual homophones. To understand whether the delayed availability of L1 semantics in L2 listening is a consequence of a general mechanism (i.e., a target-language priority) that the bilingual employs in order to avoid between-language competition or whether it rather reflects a specific feature of L2 processing it is necessary to ascertain whether the delayed availability of non-target language semantics also occurs in native-language listening. To this end, we performed two additional ERP experiments with an independent group of Dutch–English bilinguals, this time focussing on L1 speech comprehension. The results of Experiment 4 confirmed that the semantics of language-switched words seem to become available after a short delay. Interestingly, however, results from Experiment 3 suggest immediate (rather than delayed) availability of the L2 semantics of homophones in L1 contexts.

3.5.1. *Functional significance of the N400*

Our rationale underlying the design of the four experiments was that the time courses and amplitudes of negativities induced by homophones and language switches would be indicative for the semantic activation of the target words to which they occurred. There is no general consensus on the processes giving rise to the N400 (for a discussion of current theories of the N400 see Kutas and Federmeier (2011)). Nonetheless, most researchers agree that the modulation of the N400 in response to the congruency of a target word with the preceding context depends on semantic activation of the target word, either reflecting the ease of semantic integration as such (Kutas and Hillyard, 1984; Brown and Hagoort, 1993; Hagoort et al., 2009) or the ease of lexical access of the target word due to semantic constraints provided by the context (Kutas and Federmeier, 2000; Lau et al., 2008). Semantic integration accounts are sometimes referred to a 'late' or 'post-lexical' accounts, when semantic activation of the target word is assumed to occur after word recognition (Brown and Hagoort, 1993). However, semantic integration does not necessarily presuppose that the target word has been recognised. N400 effects can be elicited by non-lexical stimuli, such as pseudowords (Deacon et al., 2004) and have been observed to arise before target word recognition (Van den Brink et al., 2006; FitzPatrick and Indefrey, 2010), suggesting that semantic features of several lexical candidates can be activated before one of them is eventually selected and recognised (Zwitserlood, 1989; see also Kutas and Federmeier, 2011). Insofar as semantic features are available for integration immediately upon activation, our assumption that the N400 indexes the semantic activation of target words (more specifically: target word candidates) is compatible with both lexical access and (early) integration accounts.

A possible alternative explanation for the occurrence of a transient negativity for language switches, and *non-target biased* homophones in the L2 could be that the ascending flank of the N400 is not sensitive to a semantic but a phonological mismatch between the critical word and a form-level expectation based on the sentence context. A number of studies (Connolly and Phillips, 1994; Connolly et al., 1992, 1990; Newman et al., 2003; Van Den Brink and

Hagoort, 2004; Van Den Brink et al., 2001) find a separable ERP component on the ascending flank of the N400 (but see Diaz and Swaab, 2007; FitzPatrick and Indefrey, 2010), which is often referred to as the Phonological Mismatch Negativity (PMN). This component is purported to be involved in comparing the incoming acoustic data with a pre-activated form representation of an expected lexical item (Connolly et al., 1990, 1992; Connolly and Phillips, 1994; Newman et al., 2003). A mismatch between a predicted lexical item (the highest cloze probability item for a given context) and the target words in the *fitting switch*, and *non-target bias* conditions would thus lead to a PMN response, but should not elicit an N400 effect (due to the fact that these targets are nevertheless semantically congruent with the context). This interpretation is, however, not compatible with the observation that homophones in a *non-target biased* L1 sentence context (e.g., "Zijn kat is zijn favoriete *pet*"; EN: "His cat is his favourite *hat*") did not elicit a negative effect (compared to *semantically congruent* words) despite the fact that the Dutch sentence completion (e.g., *huisdier* 'pet') should be predicted here. We therefore consider it unlikely that the transient negativities elicited by language switches and non-target biased homophones (in L2) are due to phonological mismatches with predicted lexical items.

3.5.2. *Activation of both meanings of interlingual homophones*

Our approach of examining the semantic activation of homophones is comparable to studies on the processing of intralingual homophones and our observed pattern of the N400 responses to interlingual homophones in the *target bias* conditions and the *fully incongruent* conditions (Experiments 1 and 3) seem remarkably similar to ERP findings on intralingual homophones suggesting (a) that both contextually appropriate and contextually inappropriate homophone/homograph meanings can be activated (Swaab et al., 2003) and (b) that the activation time courses may differ (Van Petten and Kutas, 1987).

One account of the activation of both meanings of interlingual homophones could be that the bilingual language comprehension system assesses the semantic fit of both the L1 and L2 meanings of the homophone in an additive manner. Such an account would make the following predictions: (1) when neither meaning of the homophone semantically fits the context a large N400 is elicited, (2) should both meanings of the homophone semantically fit the context no N400 effect should emerge (compared to semantically congruent words), and (3) when only one of the meanings of the homophone is congruent with the context an intermediate N400 effect (attenuated compared to fully incongruent words and enhanced compared to semantically congruent words) should occur. Two reasons, however, render this 'additive semantic evaluation' account implausible. Firstly, such an explanation cannot account for the time-course differences of the negative effects for *target biased* and *non-target biased* homophones in L2 sentence contexts, which were observed in Experiment 1. Secondly, 'additive semantic evaluation' should result in intermediate N400 responses for both *target biased* and *non-target biased* homophones in L1 sentence contexts; however our data (Experiment 3) did not show significant N400 effects (compared to semantically congruent conditions) in either of these conditions.

At first glance, a more promising way to explain the sequential nature of target, and non-target language semantic activation to interlingual homophones in L2 contexts seem to be (re)ordered access accounts of ambiguous word processing (Duffy et al., 1988; e.g., Hogaboam and Perfetti, 1975). These accounts assume that access to ambiguous word meanings proceeds according to word frequency in the absence of a biasing sentence context and that a strongly biasing context can effectively reorder the sequence of word meaning activations. Although the bilinguals in our study were highly proficient, all of them had vastly more language experience in Dutch than in English (on average, 20 years versus 8 years). This would suggest that the L1 meaning of the homophone should have a subjectively higher frequency than the L2 meaning of the homophone. In essence, this would predict that the L1 meaning of the homophone should be accessed first unless the context biases towards the L2 meaning of the homophone. Our data, however, suggest that (in L2 contexts) the target language meaning of the homophone is accessed first regardless of the contextual bias. In L1 contexts, the absence of an N400 effect when either of the homophone meanings was congruent within the sentence context seems to be more in line with exhaustive access accounts (e.g., Swinney, 1979). Our data are, thus, difficult to account for by merely extending models of monolingual ambiguous word processing.

3.5.3. *Target language head start*

Inspection of the ERP waveforms for homophones and language switches (Figs. 1 and 2) reveals an interesting parallel which might help us arrive at an explanation for the observed time-course differences in between-language homophone meaning activation. Particularly, the fact that language switches (irrespective of the context language) also elicited a transient negative effect which offset earlier than the N400 effect observed between *fully incongruent* words (both switched, and non-switched) and *fully congruent* words, could be argued to have a similar processing origin as the transient negative effect observed for *non-target biased* homophones in L2 sentence contexts. The most parsimonious account for these data is that the transient negative effects reflect the initial unavailability of between language semantics (for a similar account for N400 effects in language switching in visual word recognition see Van Der Meij et al. (2010)). That is, in the case of semantically congruent language switches, the language-switched word is initially perceived as incongruent with the context. This leads to an initial negative effect, which is aborted once the semantic features of the language-switched word are retrieved (and the semantic congruity becomes apparent to the listener). The interlingual homophones activate both their L1 and their L2 meaning during speech comprehension. In L1 contexts the access to within and between language meanings occurs simultaneously. As a consequence, the semantic congruity of either the within (in the *target biased* condition) or the between language meaning (in the *non-target biased* condition) renders semantic integration easy. This is evidenced by the absence of an N400 effect for *target biased* and *non-target* biased homophones (compared to *fully congruent* words). In L2 contexts the target language (i.e., L2) meaning of the homophone is accessed first followed by the non-target language (i.e., L1) meaning. This sequential meaning activation results in a transient negative effect (compared to *fully congruent* words) for *non-target biased* homophones (initially the semantically incongruent target language meaning is activated, followed by the semantically congruent non-target language meaning) and a later negative effect for *target biased* homophones (initially the semantically congruent target language meaning is activated followed by the semantically incongruent non-target language meaning).

3.5.4. *Asymmetry of target language priority for homophones in L1 and L2 listening*

If our interpretation of transient negative effects reflecting the initial unavailability of between language semantics is correct, we have to assume that homophone meanings seem to be activated simultaneously in L1 contexts and sequentially (target language followed by non-target language) in L2 contexts.[2] One way to account for this finding is to draw a parallel with similar asymmetries, which have been observed in single-word studies of language switching in both comprehension and production. In the domain of language production, it seems that when bilinguals (at least unbalanced bilinguals; see also Costa and Santesteban (2004)) are tasked to switch between their languages, a switch into the L1 seems disproportionally more difficult than a switch into the L2 (e.g., Meuter and Allport, 1999). In the comprehension domain the asymmetry is in the opposite direction (e.g., Grainger and Beauvillain, 1987; Von Studnitz and Green, 1997; Thomas and Allport, 2000), namely switches into the L2 seem to incur greater processing costs than switches into the L1. The most common explanation for this is that the L1 has a greater resting level of activation (or a more robust representation) than the L2 and thus (1) causes more between-language interference in language comprehension, and (2) needs to be more strongly inhibited than the L2 during language production. With respect to the present study, such an account might predict that between-language lexical candidates are more available in L2 processing than they are in L1 processing. However, our data from Experiments 1 and 3 exhibit the opposite pattern, namely both homophone meanings seem to be retrieved simultaneously during L1 comprehension, while initially only the target language meaning is retrieved during L2 comprehension. Moreover, it is important to note that the observed asymmetry does not seem to hold in the case of language switches. That is, when bilinguals in the present study heard language switches from L2 to L1 (Experiment 2) their ERPs exhibited the same transient negative effect (suggesting temporary unavailability of non-target language lexical candidates) as they did when they heard language switches from L1 to L2 (Experiment 4). In other words, retrieval of between-language semantics seems to incur a similar processing delay in both L1 listening and in L2 listening, which (if anything) is more suggestive of symmetrical switch costs.

A second possible explanation for the discrepant pattern of homophone meaning activation has to do with how we conceive of the homophones' form level representations. It could be argued that, when Dutch native speakers produce a Dutch–English homophone in English they will often employ Dutch phonological representations, in other words they will produce Dutch-accented English. Furthermore, they are also frequently exposed to the same homophone produced in a similar manner (i.e., in Dutch accented English) by other Dutch native speakers. Consequently, the mapping between the Dutch pronounced version of the homophone and its English (non-target language) meaning would presumably be stronger than the link between the English pronounced version of the homophone and its Dutch meaning (as this will arguably only be produced in this way by English-native Dutch speakers, of whom there are relatively few). In essence, stronger links between the Dutch-pronounced homophone and its between-language (English) meaning than between the English-pronounced homophone and its between-language (Dutch) meaning, imply that the Dutch-pronounced version of the homophone is 'more homophonic' than the English-pronounced version of the homophone.

To our knowledge there are no empirical studies that have directly investigated interlingual homophony in this manner; however there is some evidence from L2 processing that bilinguals listening in a non-native language accept certain mispronunciations of words while they are more unforgiving of mispronunciations in their native language. For example, Dutch–English bilinguals but not native language listeners have been shown to accept the non-word *groof* as a good exemplar for *groove* (Broersma and Cutler, 2008, 2011), *ekt* as a good exemplar for *act* (Weber et al., 2011), and either *seft* or *teft* as good exemplars for *theft* (Hanulíková and Weber, 2011) depending on their experience with foreign accented speech. In terms of the present study these findings suggest that when the bilingual hears "The policeman wore a /p^{h}Et/" the Dutch meaning of *pet* is initially unavailable (but becomes available after a short delay), because the more aspirated English pronunciation /p^{h}Et/ is a bad exemplar for the Dutch /pEt/. However, when the bilingual hears "Mijn kat is mijn

[2]The fact that bilingual lexical activation has been shown to be sensitive to language context manipulations (e.g., Elston-Guettler et al., 2005) might be a source of concern for the interpretation of our findings, given the order of participation in our experiments (counterbalanced for Experiments 1 and 2, sequential for Experiments 3 and 4). It could, for instance, be the case that for half of the participants in Experiment 1 the non-target language (L1) was unintentionally more salient due to their prior exposure to language switches in Experiment 2. In contrast, participants in Experiment 3 had not been previously exposed to the language switches (Experiment 4) and should, therefore, not have an unintentionally salient non-target language. In other words, this account would predict that the amount of between language lexical activation should be greater in Experiment 1 than in Experiment 3. Strictly speaking, we cannot rule out additional affects of experimental order on the results obtained in Experiment 1. However, the fact that the pattern of results we obtained is the opposite of this scenario (we found more cross-linguistic activation in Experiment 3 than in Experiment 1), together with the fact that an unintentionally salient non-target language does not account for the time-course differences observed in Experiment 1, renders this account unlikely. Further, an additional ANOVA on the data from Experiment 1, which included *Order of experiments* (2 levels; homophones first, language switches first) as an additional between-subjects factor, yielded no significant order effects (all $F<2$, $p=n.s.$).

lievelings /pEt/" ("My cat is my favourite *pet*") both meanings of pet are activated, because /pEt/ is considered by the bilingual to be a good exemplar of /p^{h}Et/. In other words, this means that /pEt/ could be considered 'more homophonic' than /p^{h}Et/. Obviously, this proposal challenges our conception of interlingual homophony: 'True' homophones might be exceedingly rare in L1 processing and more abundant in L2 processing, particularly in the case of L1-accented L2 speech.

3.5.5. *Functional significance of target language priority*

In the preceding sections we have interpreted our data from three of the four experiments as reflecting the activation of both target and non-target language word meanings but with a priority for the target language. We have also discussed why the activation of the English meaning of Dutch–English homophones in Dutch sentence contexts (Experiment 3) might be exceptionally fast. But if one accepts the presence of a target language priority in bilingual listeners why does it arise?

Our proposed account for the observed pattern of results is that the sentence (and/or broader language) context could impose a language-membership restriction on upcoming words, thereby attenuating the amount of between-language lexical activation. Such an account could easily be accommodated within models of monolingual speech comprehension such as Shortlist B (Norris and McQueen, 2008). In this model, the listener is assumed to behave as an optimal Bayesian decision maker, taking into account not only the incoming speech information, but also the prior probability of hearing a particular phoneme or word when recognising sounds in a speech stream. An extension of this model to the bilingual situation would then include an adjustment of the prior probabilities of lexical items based on the language membership of previous words in the speech stream (e.g., it is more likely that a word belongs to language A if previously recognised words also belong to language A). The idea of language membership restrictions is also compatible with the Activation Threshold Hypothesis for bilingual lexical activation (Paradis, 1987, 1993), which holds that the recognition threshold of non-target language items is raised during bilingual language comprehension and production. In the comprehension domain, non-target language items thus require a greater amount of bottom up support in order to pass the higher recognition threshold (and hence, might pass the threshold later than target-language lexical items). There is also empirical support for the idea that availability of between-language lexical candidates can be modulated by the language context. Elston-Guettler et al. (2005), for instance, found that cross-linguistic homograph priming in L2 sentence comprehension only occurred in participants that had previously been exposed to an L1 narrated silent film and only in the first experimental block. This led them to posit that bilinguals restrict their lexical search by gradually "zooming in" to the language at hand. This process could also be easily understood as the gradual raising of the recognition threshold of non-target language lexical items. Along these lines, the observed pattern of results in the present study could originate from a bilingual language comprehension system, which is already "zoomed in" to the target language due (in part) to our precaution of only addressing the participant in the target language prior to the experimental sessions[3].

Our account goes beyond earlier observations that between-language lexical activation is modulated by lexical and/or semantic restrictions (e.g., Altarriba et al., 1996) that are imposed by the sentence context (cf., Schwanenflugel and LaCount, 1988). Chambers and Cooke (2009), for instance, observed reduced (but not absent) between-language lexical competition, when sentence contexts were constrained towards the target language interpretation of an interlingual homophone, compared to sentences where the context provided no disambiguating information. Importantly, there are also a number of studies that show that cognate facilitation effects are reduced in semantically constraining sentences (Duyck et al., 2007; Schwartz and Kroll, 2006; Van Assche et al., 2010; Van Hell and De Groot, 2008). The common interpretation of these effects is that contextual (lexical and/or semantic) constraints reduce the number of lexical competitors that need to be considered. In a sense this would mean that in low constraint sentences, when multiple lexical candidates compete for recognition, cognates undergo facilitation due to their more shared form-level representation (e.g., Sánchez-Casas and García-Albea, 2005; Sánchez-Casas et al., 1992) or multiple activation of the same semantic features (e.g., De Groot and Nas, 1991; Kroll and Stewart, 1994). In highly constraining sentences both cognates and non-cognates experience little competition. In such a situation the added benefit of the cognates is minimal (ceiling effect). Turning to the present data, neither lexical nor semantic restrictions seem to be sufficient to account for the observed pattern of results. As outlined above (see our discussion of the phonological mismatch account for the ascending flank of the N400), the lexical restrictions view is unlikely to be able to account for the absence of incongruity effects to both *target* and *non-target biased* homophones in L1 sentence contexts (unless one assumes that two form level predictions are entertained concurrently). A semantic restrictions account would, at first glance, seem more promising. The sentence context would presumably pre-activate semantic features corresponding to one lexical entry of a homophone, thereby increasing its competitive strength with respect to its non-target language meaning. Such a mechanism would make *target biased* homophones (which match the predicted semantic features) easier to semantically integrate, while rendering *non-target biased* homophones (which do not match the predicted semantic features) more difficult to semantically integrate. However, semantic restrictions alone cannot account for the time course differences in between-language homophone meaning activation in the L2. Nor can such an account explain why semantically congruent language switches (which should match the pre-activated semantic features) are nevertheless initially difficult to

[3]Although we took precautions to keep the language environment consistent for each of our experiments (addressing the participant in only the 'context' language of the experiment), this manipulation was unlikely to have put the participants in a completely 'monolingual language mode'. We can, thus, not exclude that between-language lexical activation could have been absent if we had successfully managed to put our participants in a monolingual language mode.

semantically integrate. To account for the target language priority, we argue that it is necessary to assume that the (language and/or sentence) context imposes language-membership restrictions on upcoming words in the speech stream in addition to semantic restrictions. These language membership restrictions render non-target language lexical candidates unavailable for semantic integration (e.g., by virtue of having lower prior probabilities or having to pass a higher recognition threshold) until they have garnered enough bottom-up support from the speech stream.

In summary, the present study shows that the bilingual cannot help but activate both meanings of cross-linguistic homophones, irrespective of whether they make sense within the current context or not. Homophone meanings seem to be accessed simultaneously in L1 contexts, but exhibit a target-language priority in L2 contexts. The same target language priority is observed for language switches in both L1 and L2 contexts. A head-start for the target language may help attenuate adverse effects of between-language lexical competition.

4. Experimental procedure

4.1. Experiments 1 & 2

4.1.1. Participants

Thirty right-handed Dutch–English bilinguals (5 males; average age: 21.3 years) participated in the experiment. Data from five participants were excluded from further analysis because too few trials remained after artefact rejection, leaving 25 participants in the final analysis. The participants' English proficiency was assessed using 50 grammaticality judgement items of the Oxford Placement Test (Allan, 1992) (mean score: 43.03, "advanced level", $SD=4.05$; maximum score: 50) and a non-speeded lexical decision test (60 items), created by Meara (1996) and later adapted by Lemhöfer et al. (2004)(mean $\Delta M=0.47$, $SD=0.29$; maximum score: 1.0). Additionally participants completed an extensive language history questionnaire, which included a self-reported English proficiency measure (mean score: 3.9 out of 5). Participants were either paid a small fee or they received study credits. None of the participants had any neurological impairment. All participants gave their written informed consent.

4.1.2. Materials

Aside from the language-switched items in Experiment 2 (see below), all experimental materials were in the subjects' second language (English). The experimental sentences, fillers, and practice items were spoken by one of the authors (a male, English–Dutch balanced bilingual) with normal intonation and at a normal speaking rate. The materials were digitally recorded in a sound-attenuating booth and digitised at a rate of 44.1 kHz. Sound files were later equalised to eliminate any differences in sound level.

For Experiment 1 we chose 56 monosyllabic English–Dutch non-cognate interlingual homophones (e.g., *pet*, Dutch: hat) based on their phonetic transcriptions. For each of the homophones we created three sentence frames. The first sentence frame was semantically congruent with the English meaning of the homophone (*target bias* condition; e.g., "My cat is my favourite *pet* in the world."), the second was semantically congruent with the Dutch meaning of the homophone (*non-target bias* condition; e.g., "The policeman wore a *pet* on his head."), and the third was semantically incongruent both in English and in Dutch (*fully incongruent* condition; e.g., "Jeremy drove the *pet* on a racetrack."). We created one further sentence frame paired with a non-homophonic semantically congruent critical word (*fully congruent* condition; e.g., "We went to the Vatican to see the *Pope* give an address"). We thus ended up with four conditions, each with 56 items. We additionally included 112 semantically congruent filler sentences. All sentence frames were cloze tested by an independent group of Dutch–English bilinguals. The cloze probabilities for the critical words used in Experiment 1 were 0.72 in the *fully congruent* condition, 0.00 in the *fully incongruent* condition, 0.60 in the *target bias* condition, and 0.00 in the *non-target bias* condition. Participants in the cloze test were also asked to provide a Dutch completion for the sentences in the *non-target bias* condition. Participants provided the intended interlingual homophone in 56% of these sentences (i.e., cloze probability$=0.56$). Across both Experiment 1 and Experiment 2 each condition was controlled with respect to cloze probability (English and Dutch, where applicable) and word frequency (average frequency per million: *fully congruent* condition$=40.14$, English homophone meaning$=40.47$, Dutch homophone meaning$=34.34$). Word frequencies were taken from the CELEX lemma database (Baayen et al., 1993). Within each sentence frame the critical word was either the direct or indirect object of the verb and across conditions all sentence materials preceding the critical word were of approximately equal length. To ensure experimental data were not influenced by end-of-sentence wrap-up effects critical words never occurred in a sentence-final position. Due to design constraints each participant heard each interlingual homophone in three different sentence contexts. To minimise potential order effects we created six pseudo-randomised stimulus lists across which the order in which each homophone token occurred was balanced. Each stimulus list was presented to an equal number of participants.

To make certain that the homophones were not unintentionally uttered with a more Dutch sounding accent in any of the conditions, an independent group of participants ($n=10$) judged the 3 tokens of each homophone (56 homophones $\times$ 3 conditions) with respect to whether they could perceive any trace of a Dutch accent. The homophones were spliced out of the original stimulus materials and were presented in isolation. Participants heard each homophone token once and made a non-speeded judgement on a five-point scale (1=weak accent, 5=strong accent). 'Weak accent' was intentionally chosen as the lowest point on the scale in order to bias participants towards hearing an accent. A Kruskal–Wallis Test (Kruskal and Wallis, 1952) on each homophone utterance ranked by the sum of judgement scores across all participants, revealed no significant differences per condition (fully incongruent condition: $Mdn=2$, Range=1–5; non-target bias condition: $Mdn=2$, Range=1–5; target bias condition: $Mdn=2$, Range=1–5; $H(2)=0.556$, $p>0.05$).

For Experiment 2 we created four sentence frames for each of 224 monosyllabic non-homophonic critical words

(4 experimental conditions × 56 items per condition; average cloze probability for English items: *fully congruent*=0.74, *fully incongruent*=0.00; average cloze probability for Dutch items: *fitting switch*=0.79, *non-fitting switch*=0.00; average frequency per million: *fully congruent*=51.04; *fully incongruent*=44.12; *fitting switch*=69.16; *non-fitting switch*=47.50). These four frames were split over four stimulus lists such that each critical word occurred only once per list. Each stimulus list was presented to an equal number of participants. Critical words were either (a) congruent with the sentence context (*fully congruent* condition), (b) incongruent with the sentence context (*fully incongruent* condition), (c) L1 lexical items that were congruent with the sentence context (*fitting switch* condition), or (d) L1 lexical items that were incongruent with the sentence context (*non-fitting switch* condition). We additionally added 112 semantically congruent filler sentences to each stimulus list.

4.1.3. Procedure

Half the participants first participated in Experiment 1 the other half first participated in Experiment 2. Participants were exclusively addressed in English by an English native speaker, both preceding and during the experiment, in order to encourage them to adopt a monolingual L2 language mode (Grosjean, 1982). Participants were placed in a sound-attenuating booth and were instructed to listen attentively to the sentences, which were played over two loudspeakers at a distance of roughly 1.5 m, and to try to understand them. The sound level was kept constant over participants. Each trial began with a 300 ms warning tone, followed by 1200 ms of silence, then a spoken sentence. The next trial began 4100 ms after the sentence offset. To ensure that participants did not blink during and shortly after presentation of the sentence, 1000 ms prior to the beginning of the sentence a fixation point was displayed. Participants were instructed not to blink while the fixation point was on the screen. The fixation point remained until 1600 ms after the offset of the spoken sentence.

Prior to the experiment the participants heard 10 practice sentences, half of which were semantically congruent and half included a semantically incongruent word. The experimental session was split into 3 blocks of approximately equal duration. In between blocks participants had an optional 5–10 min break. To ensure that participants remained focused on the task, they were prompted to make an animacy decision regarding the previous sentence on three occasions per block (e.g., 'Did the previous sentence contain a living entity?'). They could respond by means of a button press. On average participants answered 93% of these questions correctly in Experiment 1 and 85% in Experiment 2, suggesting that they had attentively listened to the stimulus materials. After the EEG recording the participants completed a word translation test on the critical items from Experiments 1 and 2 to verify that they were known. On average the participants scored 92% for Experiment 1 and 87% for Experiment 2.

4.1.4. EEG recording and analysis

The EEG was recorded continuously from 63 sintered Ag/AgCl electrodes, each referred to an electrode on the nose of the participant. The electrodes were mounted in an equidistant elastic cap (http://www.easycap.de). The EEG and EOG record ings were amplified with a BrainAmp DC amplifier (Brain Products, München, Germany) using a high-cutoff of 200 Hz, a time constant of 10 s (0.016 Hz), and a sampling rate of 500 Hz. Impedances were kept below 20 kΩ with the amplifier set up for high-impedance measurement (amplifier impe dance 10 MΩ; cf., Ferree et al., 2001). Trials with deflections exceeding 70 μV were rejected. Residual blinks and eye move ments were removed from the data using a procedure based on Independent Component Analysis (ICA) as described by Jung et al. (2000).

The data were analysed using the FieldTrip (http://www.fieldtrip.fcdonders.nl) toolbox for Matlab (http://www.mathworks.com). EEG data were time-locked to critical word onset. Average waveforms were calculated for each participant using a 150 ms pre-stimulus baseline. In order to increase the signal-to-noise ratio, groups of six spatially proximate electrodes were averaged together prior to statistical analysis. The resulting nine electrode sites (see Fig. 5) were chosen in order to maintain comparability with an earlier study (FitzPatrick and Indefrey, 2010). Statistical analysis was per formed by taking the mean amplitude per site from the grand averaged data in three latency ranges (250–450 ms, 450–650 ms, and 650–850 ms relative to critical word onset) based on Moreno et al. (2002).

In Experiment 1 we used an omnibus analysis of variance (ANOVA) in each latency range with condition (4 levels) and site (9 levels) as within subject factors. In Experiment 2 we used an omnibus analysis of variance (ANOVA) in each latency range with congruity (2 levels), switching (2 levels) and site (9 levels) as within subject factors. Seven electrodes were excluded from the analysis in order to have an equal number of electrodes in each site (see Fig. 5). All *p* values are reported after Greenhouse–Geisser correction (Greenhouse and Geisser, 1959). Contrasts between pairs of conditions were tested using a cluster randomisation approach (in the latency range between 200 ms and 2000 ms) that corrects for multiple comparisons (Maris, 2004). This technique provides us with an estimate of the onset and offset of differences between conditions, but is more conservative than the repeated-measures ANOVA (for a brief description see Takashima et al. (2006); Tuladhar et al. (2007)). In Experiment 1, cluster randomisation was performed on the following pairs of conditions: *Fully incongruent* versus *fully congruent*, *non-target bias* versus *fully congruent*, and *target bias* versus *fully congruent*. In Experiment 2, cluster randomisation was performed on the following pairs of conditions: *fully incongruent* versus *fully congruent*; *fitting switch* versus *fully congruent*; *non-fitting switch* versus *fully congruent*; *non-fitting switch* versus *fully incongruent*.

4.2. Experiments 3 and 4

4.2.1. Participants

Forty proficient Dutch–English bilinguals participated in both Experiments 3 and 4 (on separate days). Two of the participants were excluded due to excessive alpha in the EEG signal, a further five were excluded due to eye blink contamination of more than 50% of experimental trials. This left 33 participants in the final analysis. On average, the participants scored 42 on the Oxford placement test ("advanced level", $SD=1.41$; maximum score: 50), obtained a mean ΔM of 0.69 ($SD=0.14$; maximum score: 1.0) on the lexical decision test, and rated their English proficiency as 3.9 out of 5. Participants were either paid a small fee or they received study credits. None of the participants had any neurological impairment. All participants gave their written informed consent.

4.2.2. *Materials*

The speaker and recording procedure for the stimulus materials were identical to Experiments 1 and 2.

In Experiment 3, all the experimental sentences and homophones were translated versions of those used in Experiment 1 (56 homophones used in 3 conditions yielding a total of 168 experimental sentences), with the exception of the fully congruent condition. We created an additional 56 Dutch sentences with 56 monosyllabic, frequency matched, semantically congruent, Dutch items for the fully congruent condition. This gave us a total of 224 experimental sentences (56 per condition; cloze probabilities: *fully congruent*=0.72, *target bias*=0.60, *non-target bias* Dutch=0.00, *non-target bias* English=0.56; average frequency per million: *fully congruent* condition=39.89, English homophone meaning=40.47, Dutch homophone meaning=34.34). As in Experiment 1, each participant heard each interlingual homophone in three different sentence contexts. To minimise potential order effects we created six pseudo-randomised stimulus lists across which the order in which each homophone token occurred was balanced. Each stimulus list was presented to an equal number of participants. We additionally added 112 semantically congruent Dutch filler sentences to each stimulus list.

To make certain that the homophones were not unintentionally uttered with a more English sounding accent in any of the conditions, an independent group of participants ($n=11$) judged 3 tokens of each homophone (56 homophones $\times$ 3 conditions) with respect to whether they could perceive any trace of an English accent. The homophones were spliced out of the original stimulus materials and were presented in isolation. Participants heard each homophone token once and made a non-speeded judgement on a five-point scale (1=weak accent, 5=strong accent). A Kruskal–Wallis Test (Kruskal and Wallis, 1952) on each homophone utterance ranked by the sum of judgement scores across all participants, revealed no significant differences per condition (target bias: $Mdn=2$, Range=1–5; non-target bias: $Mdn=2$, Range=1–5; target bias: $Mdn=2$, Range=1–5; $H(2)=0.827$, $p>0.05$).

The stimulus sentences in Experiment 4 were translated from those used in Experiment 2. We thus had 224 experimental sentences (56 per condition; average cloze probability for Dutch items: *fully congruent*=0.44, *fully incongruent*=0.00; average cloze probability for English items: *fitting switch*=0.44, *non-fitting switch*=0.00; average frequency per million: *fully congruent*=69.16; *fully incongruent*=47.50; *fitting switch*=51.04; *non-fitting switch*=44.12). The sentences were assigned to 4 pseudo-randomised stimulus lists along with 112 semantically congruent filler sentences. Each stimulus list was presented to an equal number of participants.

4.2.3. *Procedure*

The experimental procedure was identical to Experiments 1 and 2. In these experiments the participants were only addressed in Dutch (L1) in order to encourage them to adopt a monolingual L1 language mode (Grosjean, 1982). In Experiment 3, participants answered 80% of the animacy questions correctly and in Experiment 4, participants answered 74% of the animacy questions correctly, suggesting that they had attentively listened to the stimulus materials. After the EEG recording in Experiment 4 the participants completed a word translation test on the English homophone translations from Experiment 3 and the English lexical items (language switches) from Experiment 4 to verify that they were known. On average the participants translated 86% of the items correctly.

4.2.4. *EEG recording and analysis*

The EEG recording procedure was identical to Experiments 1 and 2.

Acknowledgements

The authors would like to thank Ton Dijkstra, Janet van Hell, Guillaume Thierry and two anonymous reviewers for their helpful comments and suggestions. Further thanks go to Frauke Hellwig, Henrik Thiemann, Juliane Palm, Jana Basnakova, Lilla Magyari, and Kirsten Weber for their help in applying EEG electrodes.

Appendix A. Supplementary materials

Supplementary materials associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.brainres.2013.10.014.

REFERENCES

Allan, D., 1992. Oxford Placement Test. Oxord Univeristy Press, Oxford.

Altarriba, J., Kroll, J.F., Sholl, A., Rayner, K., 1996. The influence of lexical and conceptual constraints on reading mixed-language sentences: evidence from eye fixations and naming times. Mem. Cognit. 24 (4), 477–492.

Ardal, S., Donald, M.W., Meuter, R., Muldrew, S., Luce, M., 1990. Brain responses to semantic incongruity in bilinguals. Brain Lang. 39 (2), 187–205.

Baayen, R.H., Piepenbrock, R., van Rijn, H., 1993. The CELEX Lexical Database. University of Pennsylvania, Linguistic Data Consortium, Philadelphia, PA.

Broersma, M., Cutler, A., 2008. Phantom word activation in L2. System 36 (1), 22–34.

Broersma, M., Cutler, A., 2011. Competition dynamics of second-language listening. Q. J. Exp. Psychol. 64 (1), 74–95.

Brown, C.M., Hagoort, P., 1993. The processing nature of the N400 – evidence from masked priming. J. Cogn. Neurosci. 5 (1), 34–44.

Chambers, C., Cooke, H., 2009. Lexical competition during second-language listening: sentence context, but not proficiency,

constrains interference from the native lexicon. J. Exp. Psychol. Hum. Percept. Perform. 35 (4), 1029–1040.
Connolly, J.F., Phillips, N.A., 1994. Event-related potential components reflect phonological and semantic processing of the terminal word of spoken sentences. J. Cogn. Neurosci. 6 (3), 256–266.
Connolly, J.F., Phillips, N.A., Stewart, S.H., Brake, W.G., 1992. Event-related potential sensitivity to acoustic and semantic properties of terminal words in sentences. Brain Lang. 43 (1), 1–18.
Connolly, J.F., Stewart, S.H., Phillips, N.A., 1990. The effects of processing requirements on neurophysiological responses to spoken sentences. Brain Lang. 39 (2), 302–318.
Costa, A., Santesteban, M., 2004. Bilingual word perception and production: two sides of the same coin?. Trends Cogn. Sci. 8 (6), 253.
De Groot, A.M.B., Nas, G.L.J., 1991. Lexical representation of cognates and noncognates in compound bilinguals. J. Mem. Lang. 30 (1), 90–123.
Deacon, D., Dynowska, A., Ritter, W., Grose-Fifer, J., 2004. Repetition and semantic priming of nonwords: implications for theories of N400 and word recognition. Psychophysiology 41 (1), 60–74.
Degani, T., Tokowicz, N., 2010. Semantic ambiguity within and across languages: an integrative review. Q. J. Exp. Psychol. 63 (7), 1266–1303.
Diaz, M.T., Swaab, T.Y., 2007. Electrophysiological differentiation of phonological and semantic integration in word and sentence contexts. Brain Res. 1146, 85–100.
Dijkstra, T., 2005. Bilingual visual word recognition and lexical access. Handbook of Bilingualism: Psycholinguistic Approaches, 179–201.
Dijkstra, T., van Heuven, W.J.B.V., 1998. The BIA model and bilingual word recognition. In: Grainger, J., Jacobs, A. (Eds.), Localist Connectionist Approaches to Human Cognition. Erlbaum, Mahwah, NJ, pp. 189–225.
Dijkstra, T., van Heuven, W.J.B.V., 2002. The architecture of the bilingual word recognition system: from identification to decision. Biling 5 (03), 175–197.
Duffy, S., Morris, R., Rayner, K., 1988. Lexical ambiguity and fixation times in reading. J. Mem. Lang. 27 (4), 429–446.
Duyck, W., Van Assche, E., 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), 667–679.
Elston-Guettler, K.E., Gunter, T.C., Kotz, S.A., 2005. Zooming into L2: global language context and adjustment affect processing of interlingual homographs in sentences. Brain Res. Cogn. Brain Res. 25 (1), 57–70.
Ferree, T., Luu, P., Russell, G., Tucker, D.M., 2001. Scalp electrode impedance, infection risk, and EEG data quality. Clin. Neurophysiol. 112 (3), 536–544.
FitzPatrick, I., 2011. Lexical Interactions in Non-Native Speech Comprehension: Evidence from Electro-Encephalography, Eye-Tracking, and Functional Magnetic Resonance Imaging (unpublished Doctoral dissertation). Radboud University Nijmegen, Nijmegen, The Netherlands.
FitzPatrick, I., Indefrey, P., 2010. Lexical competition in nonnative speech comprehension. J. Cogn. Neurosci. 22 (6), 1165–1178.
Glucksberg, S., Kreuz, R.J., Rho, S.H., 1986. Context can constrain lexical access: implications for models of language comprehension. J. Exp. Psychol. Learn. Mem. Cogn. 12 (3), 323–335.
Grainger, J., Beauvillain, C., 1987. Language blocking and lexical access in bilinguals. Q. J. Exp. Psychol. A 39 (2), 295–319.
Grainger, J., Dijkstra, T., 1992. On the representation and use of language information in bilinguals. In: Harris, R.J. (Ed.), Cognitive Processing in Bilinguals. Elsevier Science Publishers, Amsterdam, pp. 207–220.
Greenhouse, S., Geisser, S., 1959. On methods in the analysis of profile data. Psychometrika 24 (2), 95–112.
Grosjean, F., 1982. Life with Two Languages: An Introduction to Bilingualism. Harvard University Press, Cambridge, MA.
Hagoort, P., Baggio, G., Willems, R.M., 2009. Semantic unification. In: Gazzaniga, M.S. (Ed.), The Cognitive Neurosciences 4th ed. The MIT Press, Cambridge, MA, pp. 819–836.
Hagoort, P., Brown, C.M., 2000. ERP effects of listening to speech: semantic ERP effects. Neuropsychologia 38 (11), 1518–1530.
Hahne, A., 2001. What's different in second-language processing? Evidence from event-related brain potentials. J. Psycholinguist. Res. 30 (3), 251–266.
Hanulíková, A., Weber, A., 2011. Sink positive: linguistic experience with th substitutions influences nonnative word recognition. Atten. Percept. Psychophys. 74 (3), 613–629.
Hogaboam, T., Perfetti, C.A., 1975. Lexical ambiguity and sentence comprehension. J. Verbal Learn. Verbal Behav. 14 (3), 265–274.
Johnson, R., 1986. A triarchic model of P300 amplitude. Psychophysiology 23 (4), 367–384.
Ju, M., Luce, P.A., 2004. Falling on sensitive ears: constraints on bilingual lexical activation. Psychol. Sci. 15 (5), 314–318.
Jung, T.-P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T.J., 2000. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin. Neurophysiol. 111 (10), 1745–1758.
Kroll, J.F., Stewart, E., 1994. Category interference in translation and picture naming: evidence for asymmetric connections between bilingual memory representations. J. Mem. Lang. 33, 149–174.
Kruskal, W., Wallis, W., 1952. Use of ranks in one-criterion variance analysis. J. Am. Stat. Assoc. 47 (260), 583–621.
Kutas, M., Federmeier, K.D., 2011. Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). Annu. Rev. Psychol. 62, 621–647.
Kutas, M., Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. Nature 307 (5947), 161–163.
Kutas, M., Federmeier, K.D., 2000. Electrophysiology reveals semantic memory use in language comprehension. Trends in Cognitive Sciences 4 (12), 463–470.
Lagrou, E., Hartsuiker, R.J., Duyck, W., 2011. Knowledge of a second language influences auditory word recognition in the native language. J. Exp. Psychol. Hum. Percept. Perform. 37 (4), 952.
Lau, E.F., Phillips, C., Poeppel, D., 2008. A cortical network for semantics: (de) constructing the N400. Nature Reviews Neuroscience 9 (12), 920–933, http://dx.doi.org/10.1038/nrn2532.
Lemhöfer, K., Dijkstra, T., Michel, M., 2004. Three languages, one ECHO: cognate efects in trilingual word recognition. Lang. Cognit. Process. 19 (5), 585–611.
Li, P., Yip, M.C., 1998. Context effects and the processing of spoken homophones. Read. Writ. 10 (3–5), 223–243.
Libben, M.R., Titone, D.A., 2009. Bilingual lexical access in context: evidence from eye movements during reading. J. Exp. Psychol. Hum. Percept. Perform. 35 (2), 381–390.
Marian, V., Spivey, M., 2003a. Competing activation in bilingual language processing: within- and between-language competition. Biling 6 (2), 97–115.
Marian, V., Spivey, M., 2003b. Bilingual and monolingual processing of competing lexical items. Appl. Psycholinguist. 24 (2), 173–193.
Marian, V., Spivey, M., Hirsch, J., 2003. Shared and separate systems in bilingual language processing: converging

evidence from eyetracking and brain imaging. Brain Lang. 86 (1), 70–82.
Maris, E., 2004. Randomisation tests for ERP topographies and whole spatiotemporal data matrices. Psychophysiology 41 (1), 142–151.
Martin, C.D., Dering, B., Thomas, E.M., Thierry, G., 2009. Brain potentials reveal semantic priming in both the 'active' and the 'non-attended' language of early bilinguals. Neuroimage 47 (1), 1–8.
Meara, P., 1996. English Vocabulary Tests: 10K. Centre for Applied Language Studies, Swansea, UK.
Meuter, R., Allport, A., 1999. Bilingual language switching in naming: asymmetrical costs of language selection. J. Mem. Lang. 40 (1), 25–40.
Moreno, E.M., Federmeier, K.D., Kutas, M., 2002. Switching languages, switching Palabras (words): an electrophysiological study of code switching. Brain Lang. 80, 188–207.
Newman, R.L., Connolly, J.F., Service, E., McIvor, K., 2003. Influence of phonological expectations during a phoneme deletion task: evidence from event-related brain potentials. Psychophysiology 40 (4), 640–647.
Norris, D., McQueen, J.M., 2008. Shortlist B: a Bayesian model of continuous speech recognition. Psychol. Rev. 115 (2), 45.
O'Rourke, T.B., Holcomb, P.J., 2002. Electrophysiological evidence for the efficiency of spoken word processing. Biol. Psychol. 60 (2–3), 121–150.
Paradis, M., 1987. Neurolinguistic perspectives on bilingualism. In: Paradis, M., Libben, G. (Eds.), The Assessment of Bilingual Aphasia. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ, pp. 1–17.
Paradis, M., 1993. Linguistic, psycholinguistic, and neurolinguistic aspects of "interference" in bilingual speakers: the activation threshold hypothesis. Int. J. Psycholinguist. 9, 133–145.
Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. 118 (10), 2128–2148.
Praamstra, P., Meyer, A.S., Levelt, W.J.M., 1994. Neurophysiological manifestations of phonological processing - latency variation of a negative Erp component timelocked to phonological mismatch. J. Cogn. Neurosci. 6 (3), 204–219.
Rayner, K., 1998. Eye movements in reading and information processing: 20 years of research. Psychol. Bull. 124 (3), 372–422.
Sánchez-Casas, R.M., García-Albea, J., 2005. The representation of cognate and noncognate words in bilingual memory, Handbook of Bilingualism: Psycholinguistic Approaches226–250.
Sánchez-Casas, R.M., García-Albea, J.E., Davis, C.W., 1992. Bilingual lexical processing: exploring the cognate/non-cognate distinction. Eur. J. Cogn. Psychol. 4 (4), 293–310.
Schulpen, B., Dijkstra, T., Schriefers, H.J., Hasper, M., 2003. Recognition of interlingual homophones in bilingual auditory word recognition. J. Exp. Psychol. Hum. Percept. Perform. 29 (6), 1155–1178.
Schwanenflugel, P., LaCount, K., 1988. Semantic relatedness and the scope of facilitation for upcoming words in sentences. J. Exp. Psychol. Learn. Mem. Cogn. 14 (2), 344–354.
Schwartz, A., Arêas Da Luz Fontes, A.B., 2008. Cross-language mediated priming: effects of context and lexical relationship. Biling 11 (1), 95–110.
Schwartz, A., Kroll, J.F., 2006. Bilingual lexical activation in sentence context. J. Mem. Lang. 55 (2), 197–212.
Simpson, G., 1981. Meaning dominance and semantic context in the processing of lexical ambiguity. J. Verbal Learn. Verbal Behav. 20 (1), 120–136.
Simpson, G.B., Kang, H., 1994. Inhibitory processes in the recognition of homograph meanings. In: Dagenbach, D., Carr, T.H. (Eds.), Inhibitory Processes in Attention, Memory, and Language. Academic Press, San Diego, CA, pp. 359–387.
Spivey, M., Marian, V., 1999. Cross talk between native and second languages: partial activation of an irrelevant lexicon. Psychol. Sci. 10 (3), 281–284.
Swaab, T.Y., Brown, C.M., Hagoort, P., 2003. Understanding words in sentence contexts: the time course of ambiguity resolution. Brain Lang. 86 (2), 326–343.
Swinney, D.A., 1979. Lexical access during sentence comprehension: (re)consideration of context effects. J. Verbal Learn. Verbal Behav. 18 (6), 645–659.
Takashima, A., Jensen, O., Oostenveld, R., Maris, E., van de Coevering, M., Fernández, G., 2006. Successful declarative memory formation is associated with ongoing activity during encoding in a distributed neocortical network related to working memory: a magnetoencephalography study. Neuroscience 139 (1), 291–297.
Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M., Sedivy, J.C., 1995. Integration of visual and linguistic information in spoken language comprehension. Science 268 (5217), 1632–1634.
Thomas, M.S.C., Allport, A., 2000. Language switching costs in bilingual visual word recognition. J. Mem. Lang. 43, 44–66.
Titone, D., Libben, M., Mercier, J., Whitford, V., Pivneva, I., 2011. Bilingual lexical access during L1 sentence reading: the effects of L2 knowledge, semantic constraint, and L1-L2 intermixing. J. Exp. Psychol. Learn. Mem. Cogn. 37 (6), 1412–1431.
Tuladhar, A.M., Ter Huurne, N., Schoffelen, J.-M., Maris, E., Oostenveld, R., Jensen, O., 2007. Parieto-occipital sources account for the increase in alpha activity with working memory load. Hum. Brain Mapp. 28 (8), 785–792.
Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., Hartsuiker, R.J., 2010. The influence of semantic constraints on bilingual word recognition during sentence reading. J. Mem. Lang. 64 (1), 88–107.
Van Den Brink, D., Hagoort, P., 2004. The influence of semantic and syntactic context constraints on lexical selection and integration in spoken-word comprehension as revealed by ERPs. J. Cogn. Neurosci. 16, 1068–1084.
Van Den Brink, D., Brown, C.M., Hagoort, P., 2001. Electrophysiological evidence for early contextual influences during spoken-word recognition: N200 versus N400 effects. J. Cogn. Neurosci. 13 (7), 967–985.
Van den Brink, D., Brown, C.M., Hagoort, P., 2006. The cascaded nature of lexical selection and integration in auditory sentence processing. J. Exp. Psychol. Hum. Percept. Perform. 32 (2), 364–372.
Van Der Meij, M., Cuetos, F., Carreiras, M., Barber, H.A., 2010. Electrophysiological correlates of language switching in second language learners. Psychophysiology 48 (1), 44–54.
Van Hell, J., De Groot, A., 2008. Sentence context modulates visual word recognition and translation in bilinguals. Acta Psychol. 128 (3), 431–451.
Van Hell, J., Witteman, M., 2009. The neurocognition of switching between languages: a review of electrophysiological studies. Multidisciplinary Approaches to Code Switching, 53.
Van Petten, C., Kutas, M., 1987. Ambiguous words in context: an event-related potential analysis of the time course of meaning activation. J. Mem. Lang. 26 (2), 188–208.
Von Studnitz, R.E., Green, D.W., 1997. Lexical decision and language switching. Int. J. Biling 1 (1), 3–24.
Weber, A., Cutler, A., 2004. Lexical competition in non-native spoken-word recognition. J. Mem. Lang. 50 (1), 1–25.
Weber, A., Broersma, M., Aoyagi, M., 2011. Spoken-word recognition in foreign-accented speech by L2 listeners. J. Phon. 39 (4), 479–491.
Zwitserlood, P., 1989. The locus of the effects of sentential-semantic context in spoken-word processing. Cognition 32 (1), 25–64.