

*School of Linguistics and
Applied Language Studies*

*Wellington Working Papers
in Linguistics*

Volume 20, 2008

*ISSN 1170-1978 [print]
ISSN 2230-4681 [online]*



Wellington Working Papers in Linguistics

Volume 20, 2008

edited by Derek Wallace

**School of Linguistics and Applied Language Studies
Victoria University of Wellington
P.O. Box 600
Wellington
New Zealand**

Published 2008

Printed for
Victoria University of Wellington
P.O. Box 600
Wellington
New Zealand

ISSN 1170-1978 [print]
ISSN 2230-4681 [online]

Language-Specific Cues – A Cue to Language?

Anna Piasecki and Paul Warren

Abstract

A key issue in psycholinguistic research on the nature of the coexistence of two (or more) languages in the cognitive system of a fluent bilingual speaker include the nature of lexical access (*selective vs non-selective*). In the context of the *non-selective access* view, we investigate the extent to which sub-lexical information (eg language-specific cues, such as onset capitals for German nouns) is sufficient to constrain or eliminate lexical interaction between the bilingual speaker's languages. We also consider the extent to which the use of such information is affected by priming for a specific language from a preceding sentential context. To gain insight, experimental data from English-German bilinguals representing three different proficiency levels was collected, who listened to a sentence frame in either L1 or L2, and then performed a German (L2) lexical decision task to a word presented visually immediately after the frame. Error data shows that language-specific cues have an increasingly facilitatory effect on lexical access with increasing proficiency levels. In addition, context language effects decrease with increasing proficiency level. Response time analyses, on the other hand, reveal a delay for German-biased items, ie those with onset capitalisation. We discuss these results in the context of models of bilingual language processing.

1. Introduction

In the introduction to his chapter on visual word recognition (VWR), Balota (1994: 303) noted that '[the] word is as central to psycholinguists as the cell is to biologists'. This is reflective of the fact that VWR research has been one of the central focal points of investigation in psycholinguistics, experimental psychology and, more generally, cognitive science for more than a century now. Andrews (2006) suggests three main reasons for this. Firstly, interest in VWR arose because the ability to recognise words is the baseline for literacy. Secondly, experimental designs investigating word recognition processes provide a vehicle for exploring other cognitive processes, such as memory structures and psychopathological disorders (eg aphasia). Finally, research in this area offers crucial insights into pattern recognition and memory retrieval.

Given the importance of VWR research, it is surprising that its extension to the bilingual domain has only been relatively recent (eg Nas 1983). This is even more

surprising, considering that bilingualism¹ (if not multilingualism) is the norm in most parts of the world. Given this, it would seem important to examine bilingual VWR processes. Of course, insights from such research also have implications for how to teach vocabulary, or a second language in general.

The research presented in this paper singles out one aspect of bilingual VWR for exploration: the nature of sub-lexical information. First, in section 2 we briefly summarise relevant bilingual research to date, and introduce the main issues of relevance to bilingual VWR. In sections 3 and 4 we consider the organisation of the bilingual mental lexicon and the role played by sub-lexical information in lexical access in bilinguals. More specifically, these sections discuss and attempt to measure the extent to which language-specific information can be used to speed up the processing of presented words. Moreover, we investigate the point at which such information becomes available during the word recognition process, and the level of representation of the information used (eg sub-lexical, lexical level). Finally, we present some preliminary conclusions and suggestions for further research.

1.1 A model of bilingual lexical processing

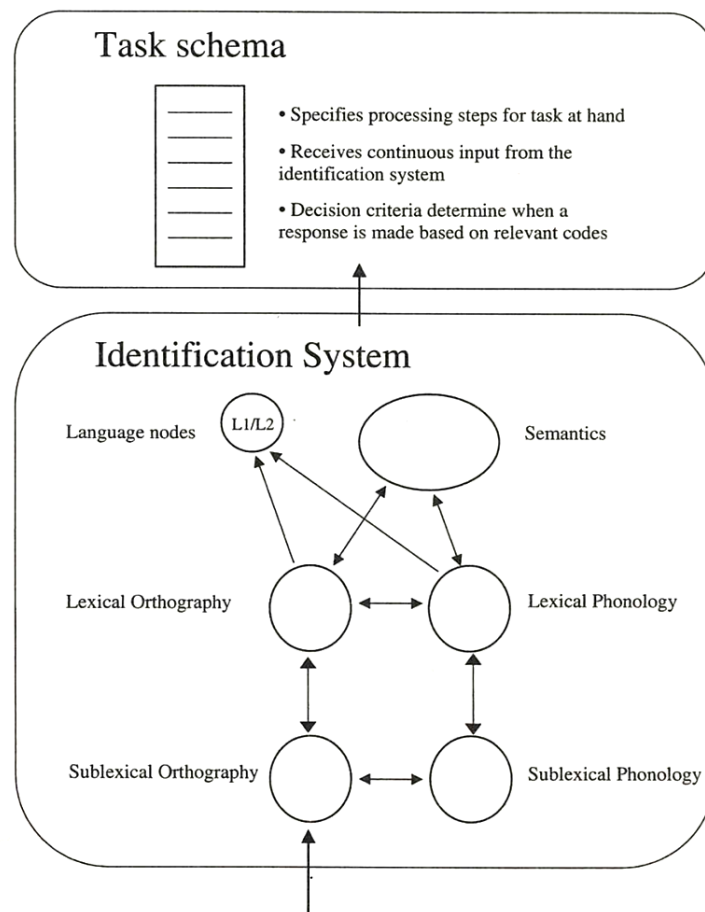
On top of the processing issues faced by a monolingual reader, bilingual readers must cope with the activation of two languages. A priori it seems reasonable to assume that two languages coexist in the cognitive system of a fluent bilingual speaker. A considerable amount of psycholinguistic research has been devoted to determining the nature of this coexistence (eg Brysbaert et al 1999; Kroll and Stewart 1994; Dijkstra and Van Heuven 1998). This research includes an increasing focus on bilingual VWR.

To date, the most prominent theoretical model of bilingual VWR, and one which provides an account for most of the recent research findings, is the Bilingual Interactive Activation Plus Model (BIA+) proposed by Dijkstra and Van Heuven (2002). This model (Figure 1) assumes that lexical information from a bilingual's two languages is represented in an integrated lexicon to which there is non-selective access (see section following for further discussion of lexical storage and access selectivity). Thus, in the initial stages of lexical retrieval, there is interactive, bottom-up, and non-selective activation of lexical information across a bilingual's languages. In terms of the model's architecture, the BIA+ contains a range of linguistic information: not only orthographic, phonological and semantic representations, but also language nodes. The orthographic and phonological representations are, in addition, extended over two processing levels, namely the sub-lexical and lexical levels. According to Dijkstra (2005: 197), access to lexical representations can be triggered solely on the basis of such linguistic information. The information flow

¹ For the sake of consistency, we have taken the term 'bilingualism' to include second and foreign language situations. Due to space limitations we will not further explore the differences between the different terms, although we are aware that the term *bilingual(ism)* is contentious. For present purposes, it is important to note that the term is being used here to include relatively high proficient second language (L2) learners.

then proceeds to the task/ decision system. The task/ decision system is assumed to be affected by extra-linguistic factors, such as participants' expectations or task demands. While these variables can in turn influence the output of the word identification system, they cannot influence the activation state of words. A further important feature of the model is the set of language nodes. These are proposed as representations of language membership.

Figure 1 The architecture of the BIA+ model (Dijkstra and Van Heuven 2002)
L1, first language; L2, second language



It has been claimed (eg Dijkstra and van Heuven 2002: 182ff.) that the BIA+ model can accommodate a large amount of research that supports non-selective access, as well as some of the more specific differences that arise across different task designs. The following section will provide a more detailed examination of some of these relevant issues (including differences between experimental tasks), and will point out areas where the model is under-elaborated (eg the relative importance of sentential context or proficiency level).

1.2 One lexicon or two?

A central issue in bilingual VWR research has been the distinction between language-dependent and language-independent lexical storage. That is, some

researchers have argued for the co-existence of two separate lexicons – one for each language – while others have argued for the existence of a single integrated lexicon for both languages.

Within research that argues for *two* mental lexicons, evidence has been presented that indicates strong cross-language connections at different levels: at the sub-lexical level (eg Brysbaert, Dyck et al. 1999), at the lexical level (eg von Studnitz and Green 2002), and/ or at the conceptual level (eg Kroll and Stewart 1994). Given those strong interfaces between languages, two questions have been addressed. Firstly, can a bilingual ever function in the L1 or L2 without constant influence of one language on the other? Secondly, how well or poorly can a bilingual activate only the appropriate language at the appropriate time and to the appropriate extent? The first issue is generally referred to as *selective* versus *non-selective access* (see discussion below). The second issue involves *cognitive control* (see Dijkstra and Van Heuven 2002 for a review of both issues).

Within research that argues for *one* mental lexicon, two further questions arise. First, when bilinguals are presented with visual stimuli, how do they know what language an input item belongs to? It is now assumed that this kind of information must be stored in the bilingual's mental lexicon for each word. Some researchers talk of a *language node* (Dijkstra and van Heuven 1998), others of a *language tag* (Green 1998). Possibly, each word has its own separate language tag/ node; alternatively, all words of one language may share their language tag/ node – more explicit information on the nature of such tags or nodes is still lacking (Dijkstra 2005). Second, if a bilingual's two languages share the same orthography or script (eg both roman script), which lexical candidate is activated (ie from L1, from L2, or from both languages) when a letter string is presented? This is a further issue that is discussed under the heading of *selective* versus *non-selective access*.

Based on evidence from a range of task designs, the majority of researchers now seem to agree that there is non-selective access of lexical information across a bilingual's two languages during VWR (eg van Hell and Dijkstra 2002; Schwartz and Kroll 2006). However, much of the crucial research has been based on the comprehension of words in isolation. In response to this, a new research direction has emerged, one which creates bilingual conditions which are more true to an everyday situation by, for instance, embedding experimental stimuli in sentential contexts. Although literature on this topic is still scarce and discussion is still at an early stage, some initial results suggest that certain factors may constrain (if not eliminate) lexical interaction between languages (Elston-Güttler, Gunter et al 2005; Duyck, Assche et al 2007). For instance, Elston-Güttler and her colleagues (2005) tested the recognition of interlingual homographs (letter strings that correspond to words in both languages) in German-English bilinguals (ie German learners of English). They used a task design in which participants had to read for comprehension a visually-presented sentence, and subsequently carry out lexical decision on a single word presented after the sentence. On critical trials, the sentences ended in an interlingual homograph (in italics in the example) and the target item for lexical decision (in small capitals in the example) was related in meaning to the non-target, L1 meaning of the homograph:

The woman gave her friend an expensive *gift* – POISON (= meaning of German word *Gift*).

When these prime-target word pairs were presented in isolation, the L2 (English) homograph always primed its L1 (German) meaning, suggesting non-selective access. However, homograph priming in sentence contexts was only found in the first three blocks of the experiment, and was absent from the remaining three blocks. Moreover, it was only found for participants who saw a German movie prior to the experiment, which increased their L1 activation. The authors' interpretation of the results was that participants adapted their lexical decision thresholds during the experimental session. They called this process 'zooming into' the all-L2 task. The authors claim that changing from one entire language context to another and staying there is likely (even in the usually less dominant L2), given a language-exclusive task. This 'adjustment of language mode settings' (Elston-Güttler et al: 58) is clearly based on Grosjean's (2001) concept of language modes (monolingual, bilingual or an in-between setting). The two concepts differ in a way that most probably reflects differences in task demands – Grosjean's concept assumes the type of continuous language-switching found in most natural bilingual situations, whereas Elston-Güttler et al refer to a complete adjustment from one monolingual setting into another monolingual setting.

A question linked to Grosjean's (2001) concept of language nodes and the findings discussed above, but one which has not received much attention, is the extent to which proficiency may have an effect on non-selective access. In one recent relevant study, Chambers and Cooke (2009) argued that context has a stronger impact than proficiency level on parallel language activation during spoken language. In their study, non-native speakers with varying proficiency levels viewed visual displays while listening to French sentences, such as:

Marie va décrire la poule (= Marie will describe the chicken).

Visual displays depicted several objects including the final noun target (eg 'chicken') and an interlingual near-homophone (eg 'pool') whose name in English is phonologically similar to the French target ('poule'). The researchers measured listeners' eye movements during target noun playback. One observation resulting from this experiment was that there was temporary lexical competition for interlingual homophones. The same pattern was reported for lower as well as higher proficiency listeners in low constraint sentences (ie sentences where there is no clear bias towards either meaning of an interlingual homophone). Apart from this finding being slightly surprising, it is possible that an entirely visual task will have a different influence on (increasingly highly proficient) bilingual lexical processing.

1.3 Language-specific sub-lexical information

A final unresolved question appears to be whether information about which language is being read or heard can be used to speed up the processing of presented words (eg Dijkstra 2003; Dijkstra 2005). To illustrate, referring to research conducted

by Kroll and Dijkstra (2002) and Schwartz et al (2000), Dijkstra (2003: 20) hypothesised that

[even] when two languages are closely related and are represented by the same script, words may contain *language-specific cues*. Examples are the diacritical markers (accents) of French and the onset capitals for nouns in German. In such cases, the use of these cues might quickly reduce the number of competitors of an item to those of the target language. [...] There is some preliminary evidence that language specific bigrams and other cues may indeed affect the selection process, but much more study is necessary here.

If language-specific information does affect the selection process, then a further question concerns the point(s) (sub-lexical level, lexical level, etc) at which such information becomes available during the word recognition process (Dijkstra 2005). A measure of the availability of such information is the extent to which it facilitates word recognition. In other words, assuming that such information is available soon enough, it might help to speed up word recognition by excluding lexical candidates from the non-target language.

2. Experiment

Given the rationale above, the aim of this study is to explore the nature of sub-lexical information (ie in the form of language-specific cues) on bilingual visual processing. To achieve this, the following research questions were addressed:

- (i) To what extent can sub-lexical information (eg in the form of language-specific cues, such as onset capitals for German nouns) facilitate or inhibit bilingual VWR? Is this information sufficient to constrain (if not eliminate) lexical interaction between the bilingual's languages?
- (ii) If sentence context affects the speed of word recognition, then bilinguals might be slower to recognise a stimulus in a language that differs from the language of the context sentence (Dijkstra 2005). Consequently, how well can a bilingual either use or discard sub-lexical information in specific language contexts?
- (iii) What effects might L2 proficiency have on the manifestation of facilitatory versus inhibitory dynamics?

2.1 Participants

Sixty-five native speakers of New Zealand English completed two experimental sessions which were approximately seven days apart. Recruited participants were current and former students from Victoria University of Wellington, with varying knowledge of German. The participants were selected to represent one of three levels of proficiency (labelled 100-, 200- and 300-level, based in most cases on their course enrolments). To test their German knowledge individually and in order to

acquire other relevant information, each participant filled out a language questionnaire and completed a German language proficiency test (adapted from Lemhöfer 2004) following the second experimental session. All participants signed a written informed consent, had normal or corrected-to-normal vision and no hearing impairment. Participants received a voucher for their participation.

2.2 Materials

During an experimental session, participants listened to a sentence fragment in either their first language (English) or their second language (German), and then performed a German lexical decision task to a word presented visually immediately after the fragment, ie they indicated whether or not the word was a real German word, by pressing one of two response buttons. As this experimental design involved an acoustic prime followed by a visual lexical decision task, primes and target items needed to be carefully selected and prepared for use. This included selecting critical target words (interlingual homographs or IHGs), selecting matched control words and nonwords for comparative analyses with critical stimuli, and then designing sentence frames (ie primes) to place these items into.

2.2.1 Selecting target words

Item construction was done in the following way. First, a list of interlingual homographs was created which was partly based on Elston-Güttler et al's (2005) item list and partly extracted from an English learner's dictionary (1999). To ensure that lower proficiency learners of German would be familiar with these items, the existing selection was matched against an entry in the vocabulary list from an elementary German learner's course book (Perlmann-Balme and Kiefer 2002) provided by a German course instructor. Meeting this criterion left us with 39 items, all of which had one meaning in English (cf *hose* = 'pipe') and another one in German (cf *hose* = 'trousers'). The Appendix contains a complete list of the 39 target words. The majority of the selected items were nouns in both languages. In a few cases, however, a German noun would belong to a different word class in English, and vice versa, or an item would belong to a different word class than a noun in both languages (commonly being an adjective, verb or adverb; usually varying across the two languages).

Note that each critical item (IHG) was presented twice in the course of the experiment, in different sessions (see further information on the experimental design below). To provide real word controls, for each IHG a pair of real word (RW) German items (eg *mut* and *uhr*) was selected using the WordGen programme (Duyck et al 2004), which uses the CELEX database (Baayen et al 1993) as a resource. One member of each RW control pair appeared in each session. These control items, consisting of 78 items in total, were matched with the set of IHGs for number of letters, number of German noun neighbours, and German log frequency per million (see Table 1). To match the critical stimulus set as closely as possible, the control RWs were mainly nouns, but also included verbs, adjectives and adverbs.

Finally, 156 nonsense words (NW) were created, again using WordGen (Duyck et al 2004) and CELEX (Baayen et al 1993). As with the RWs, pairs of NWs (78 NWs in total) were developed as matches to the IHGs, based on number of letters, number of German noun neighbours, and German bigram frequency (see Table 1). The remaining 78 NWs were matched in the same way to the set of 78 German control words (RWs). Care was taken to ensure that all nonsense words obeyed German orthographic rules and were not existing English words. Overall, half of the stimuli in each session were real words (either IHG or RW stimuli) and half were nonsense words, meaning that half of the lexical decision responses were targeted at a ‘yes’ response and other half at a ‘no’ response.

Table 1
Mean letter length, count of German noun neighbours, and frequency of different target item types, with standard deviations in parentheses

Target Items	Target letter length	Target noun neighbour count	Target frequency ^a	Target frequency ^b	
Interlingual Homographs (IHG) (N=39)	4.48 (1.04)	4.18 (2.44)	1.51 (0.78)	15035 (10669)	
Real Word Fillers (RW) (N=78)	4.48 (1.04)	4.18 (2.41)	1.50 (0.72)	14037 (10247)	
			1.50 (0.83)	12725 (10575)	
Nonsense Words (NW) (N=156)	4.48 (1.04)	4.18 (2.41)	N/ A	13017 (9194)	14410 (10136)
				14735 (10359)	14821 (10498)

^a Mean frequency per million of test and corresponding control targets, using the German log frequency in the CELEX database (Baayen et al 1993).

^b Mean frequency per million of test and corresponding control targets, using the German bigram frequency in the CELEX database (Baayen et al 1993)

2.2.2 Sentence frames

With stimulus selection completed, two sets of English sentence frames were created for each IHG, one for each control RW and one for each NW. This gave a total of 312 English sentence frames. Each English sentence frame was then translated into German, resulting in 312 German sentence frames. All sentence frames had a relatively open context with no obvious bias towards the target word meaning. With respect to the critical IHG stimuli, this means that there was no bias towards either (English or German) meaning. Finally, all English sentence frames were recorded by a native speaker of New Zealand English, and all German sentences were recorded by a native speaker of German. Two presentation lists were constructed, each containing all 312 target words or nonwords. In each presentation list half the sentence frames were in English, and half in German, rotated across lists so that if a target IHG, RW or NW was preceded by an English sentence in one list then it was

preceded by a German sentence in the other list. Within each presentation list, the targets with English and German context sentences were presented in separate sublists (ie a block of 156 English sentences and a block of 156 German sentences). These sublists were presented in separate experimental sessions one week apart. Half of the participants were exposed to the sublist with the English context sentences in the first week, and to the sublist with the German context sentences in the second week. This order of sublists was reversed for the other participants. This ensured that participants never heard the same sentence in both languages in a single session, or twice in the same language across the two sessions (see Table 2 below).

The two sublists presented to any participant included the same set of 39 critical IHG words. Each sublist had a different set of 39 matching control RWs, and a different (but matching) set of 78 NWs. All sentence frames included were unique across the two sublists (except that for any one sentence frame there was a translation equivalent of that sentence frame in the other sublist). The stimuli in each sublist were divided in six blocks, each containing 26 trials. Each block ended with a memory task (explained below) which was meant to ensure that subjects paid attention to the sentences and did not exclusively focus on the lexical decision task. The order of the six blocks was kept constant because of the limitations outlined in the following paragraph.

Participants were required to attend two sessions of approximately thirty minutes each. As explained above, in the first session participants heard sentence frames in only one of the two languages (eg English; cf Table 2). During the second experimental session, they then heard sentence frames in the other language. To control for a possible language effect, half of the participants listened to English sentences in their first session and German sentences in the second session, and the other half listened to German sentences first and English sentences in the following week's session. Stimulus order within the sessions was kept constant, so that effects of sequential order within a session (eg practice or fatigue effects) would be likely to affect each language condition equally.

Participants within each proficiency level were also allocated randomly but evenly to one of two format conditions, which related to the presentation format of the visually-presented target word. The target was either entirely in lower case, or with the first letter capitalised (referred to in this paper as Title case). For example, after the spoken fragment *The woman listened to a radio show about the perfect* the target would be either *gift* (lower case) or *Gift* (title case). The target format remained constant across both sessions for each participant. Table 2 below illustrates the resulting conditions.

Table 2
Examples of stimuli materials by condition for presentation format (Title; lower), presentation order (English sentence; German sentence), and final target item

(The use of bold and normal font for the auditory sentence primes indicates the pairing of sentences across sublists: eg Sentence 1 for HOSE in English in one sublist is paired with Sentence 2 in German in the other sublist)

Item Type	Condition(s)		Prime (auditory sentence frame)	Target (final word)	
	Item Format	Sentence Language			
Interlingual Homograph	Title Case	ENG	<i>Sentence 1 and its translation, for the target HOSE</i> (= 'trousers' in German)	Hose	
		GER	Tim's shopping list included a barbecue and a Auf Timms Einkaufsliste stand Grill und		
		GER	<i>Sentence 2 and its translation, for the target HOSE</i> (= 'trousers' in German)	Hose	
		ENG	Der Arbeiter verließ das Haus ohne The worker left the house without the		
	lower case		ENG	<i>Sentence 1 and its translation, for the target GIFT</i> (= 'poison' in German)	gift
			GER	The woman listened to a radio show about the perfect Die Frau hörte im Radio eine Sendung über das perfekte	
		GER	<i>Sentence 2 and its translation, for the target GIFT</i> (= 'poison' in German)	gift	
		ENG	Er dachte an den Keller als das beste Versteck für das He thought of the cellar as the best hideout for the		
Real Word (German)	Title Case	ENG	The aunt looked in her bag for the small	Kamm	
		GER	Die Tante suchte in ihrer Handtasche nach dem kleinen		
	lower case	GER	Alexander asked his neighbour for	mehl	
		ENG	Alexander bat seinen Nachbarn um		
Nonword	Title Case	ENG	The people loved the goofy	Arin	
		GER	Alle liebten den albernen		
	lower case	GER	The examiner carefully studied the	nark	
		ENG	Der Prüfer untersuchte sorgsam die		

2.3 Procedure

Participants were first asked to read all instructions for the experiment on the computer screen. After a short practice session, the actual test began. A trial started off by presenting an empty screen for 2500ms, a time lapse which functioned as an inter-trial interval. Immediately after that participants heard a sentence fragment over the headphones. When the sentence fragment ended, a fixation cross appeared on the empty screen. After 200ms, the fixation cross was replaced by the potential sentence-final word, ie the target, in either all lower case letters or with the first letter capitalised. The participant's task was to decide as quickly as possible whether the

word presented on screen was an existing German word, and to indicate this response by pressing one of two keys (labelled Yes and No) on a button box with millisecond timing accuracy. They were timed-out after 3000ms if they had made no response, and the next trial was started. The experiment was run in E-Prime (Schneider, Eschman et al 2002; Schneider, Eschman et al 2002) on a Windows personal computer. Different response button configurations were selected depending on whether the participant was left- or right-handed, so that every participant used their dominant hand to indicate a 'Yes' response. Between trials participants rested the index finger of each hand over the response buttons.

Participants were tested individually. To keep the entire experiment as stable as possible, the same native German-speaking researcher conducted all sessions, and the procedure was exactly the same for all participants. The lexical decision task lasted no more than 25 minutes and was presented in six blocks, as described above. At the end of each block, a memory recall task was performed which included three sentences that were previously heard over the headphones and three sentences that were not heard anywhere during the experiment. Participants were presented with these sentences on screen, including their final word, and were asked to decide whether each sentence was included in the block they had just been exposed to (ie as a combination of a spoken sentence fragment and a single completing word). This was done to ensure that subjects paid attention to the sentences and did not exclusively focus on the lexical decision task.

After the second experimental session, subjects carried out a German proficiency test, filled in a language history questionnaire, and were asked to give the English meanings of the German words represented by the IHGs in the experiment (eg for 'Hose' a correct response would be 'trousers'). The entire experimental procedure, that is both sessions, was completed in approximately 60 minutes (roughly 30 minutes per session).

2.4 Data analysis and results

Prior to data analysis, two participants had to be excluded since they did not follow the given instructions, and one further participant had to be excluded due to a high overall error rate (greater than 50%). This left data from 62 participants. Further data cleaning procedures included the exclusion of three critical IHG words and four control RWs. The three IHGs were excluded because they were not known to the majority of participants. The RWs were excluded either due to participants' high error percentage on these particular items, or because they could have been read as English words. Finally, the assignment to a particular 'proficiency' (100-, 200- or 300-) level was adjusted for three participants, after taking into account the data from their responses to the questionnaire about German language exposure and experience.

The analysis below first presents overall statistical results for error rates and for response times, before exploring effects within each level of participant proficiency, motivated by the interactions involving the Proficiency Level factor.

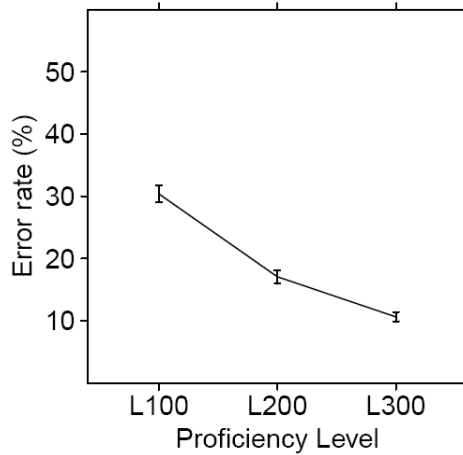
Four-way mixed effects ANOVAs were performed for the remaining data, for error rates (ie responding that a German word – either a RW or an IHG – was a nonword and vice versa) and for reaction times. The results of the ANOVA can be seen in the following set of graphs (Graphs 1–4). In the participant analysis, Sentence Language, Item Type and Proficiency Level were treated as within-participant factors, and Item Format as a between-participant factor. In the item analysis, Sentence Language, Item Format and Proficiency Level were treated as within-item factors, and Item Type as a between-item factor. Please note that although both subject and item ANOVAs were performed, only the latter will be discussed, due to space limitations.

Error rates were analysed separately for incorrect ‘nonword’ responses to real words (IHGs and RWs combined) and for incorrect ‘word’ responses to nonsense words (NWs). Both analyses revealed Proficiency Level as a strong overall effect – real words, $F(2,208) = 89.48$, $p < 0.001$; nonwords, $F(2,310) = 380.52$, $p < 0.001$ – with lower proficiency subjects making more incorrect responses to both real word and nonsense stimuli than their more proficient counterparts (see Graph 1). Proficiency Level was also involved in many interactions, such as in two-way interactions with Item Type (only possible for real words, since there is only one type of nonsense word, $F(2,208) = 26.87$, $p < 0.001$); with Item Format (real words, $F(2,208) = 3.83$, $p < 0.03$; nonwords, $F(2,310) = 42.10$, $p < 0.001$); and with Sentence Language (significant only in the nonword analysis, $F(2,310) = 4.52$, $p < 0.02$). Proficiency Level was also involved in a significant three-way interaction with Sentence Language and Item Format (real words, $F(2,208) = 6.03$, $p < 0.005$; nonwords, $F(2,310) = 3.53$, $p < 0.04$), and a marginally significant four-way interaction with Sentence Language, Item Type and Item Format (for real words only, $F(2,208) = 2.85$, $p = 0.06$).

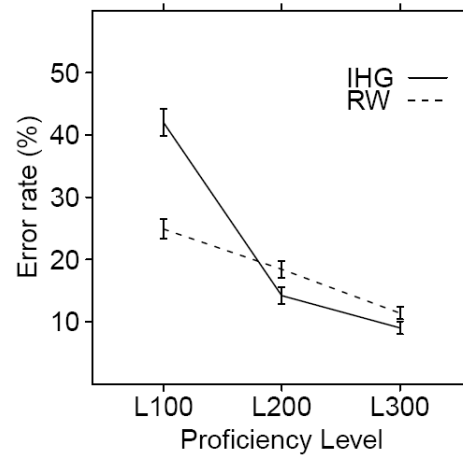
Separate response time analyses were also carried out for correct responses to real words (IHGs and RWs, including an Item Type comparison between these two) and for correct responses to nonsense words (NWs). These analyses were separated because correct responses to real words and correct responses to nonsense words involved different decision outcomes (‘yes’ and ‘no’ respectively) and required different button presses using different (dominant and non-dominant) hands. Proficiency Level again showed a strong overall effect, for both real words – $F(2,208) = 237.67$, $p < 0.001$ – and nonwords – $F(2,310) = 253.40$, $p < 0.001$ – and with increasingly faster responses to items as participants’ proficiency increased. Proficiency Level was also significantly involved in a two-way interaction with Item Format – real words, $F(2,208) = 4.60$, $p < 0.01$; nonwords, $F(2,310) = 20.36$, $p < 0.001$ – and a three-way interaction between Sentence Language and Item Format for nonwords only: $F(2,310) = 4.73$, $p < 0.009$. Response time data also revealed a strong overall effect of Item Format, for both real words – $F(1,104) = 64.95$, $p < 0.001$ – and nonwords – $F(1,155) = 175.02$, $p < 0.001$ – and its involvement in further interactions (including those already mentioned), namely with Item Type – note that this analysis is only possible for the real word contrast of IHGs and RWs: $F(1,104) = 7.43$, $p < 0.007$ – and with Sentence Language (real words, $F(1,104) = 22.08$, $p < 0.001$; nonwords, $F(1,155) = 17.69$, $p < 0.001$). The main effect of Item Type, which can only be tested for the IHG/ RW contrast in real words, was marginally significant: $F(1,104) = 3.51$, $p = 0.06$.

Graph 1

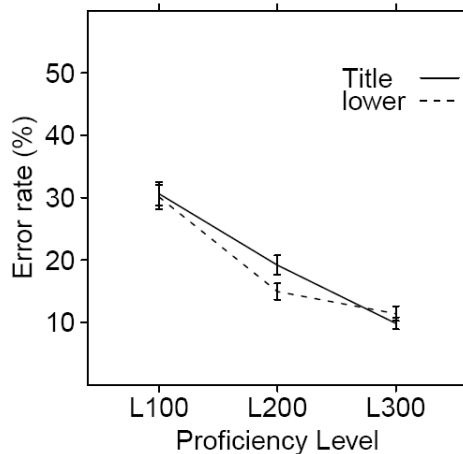
Mean incorrect responses to German real words (RWs and IHGs combined), across three proficiency levels.

**Graph 2**

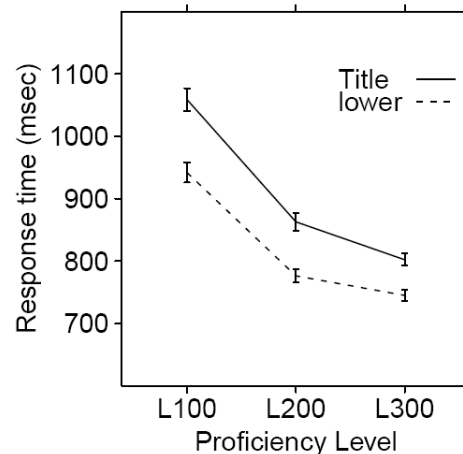
Mean incorrect responses to Interlingual Homographs vs. Real Word targets, across three proficiency levels.

**Graph 3**

Mean incorrect responses to German real word targets (RWs and IHGs combined), in Title vs. lower case, across three proficiency levels.

**Graph 4**

Mean latencies to German real word targets (IHGs and RWs combined) presented in Title vs. lower case at three different proficiency levels.



Based on the main effects and interactions found for Proficiency Level in the above analyses, separate error and RT analysis were carried out for each of the three proficiency levels (see below).

2.5 Error analysis

Recall that the overall error analysis for real words showed a main effect for Proficiency Level and an interaction of this with Item Type (RW vs IHG). Graph 2 shows that increasing proficiency results in a decrease in error rates for real words, and that this effect is greater for interlingual homographs (IHGs).

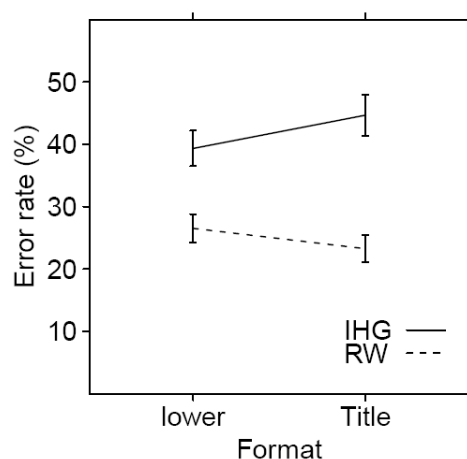
Proficiency Level also interacted with Item Format in the main analysis for both real words and nonsense words (see Graphs 3 and 4). In a separate analysis for the lower proficiency level (100) it was found that Item Format interacts with Item Type (for real words, $F(1,104) = 4.47$, $p < 0.03$), reflecting an increase in incorrect responses to

IHG with Title case (see Graph 5). Interestingly, these participants also make considerably more errors on nonwords presented in Title case ($F(1,155) = 22.30, p < 0.001$), particularly after English context sentences; the Sentence Language by Item Format was significant for errors on nonwords at 100-level: $F(1,155) = 5.23, p < 0.02$ (see Graph 6).

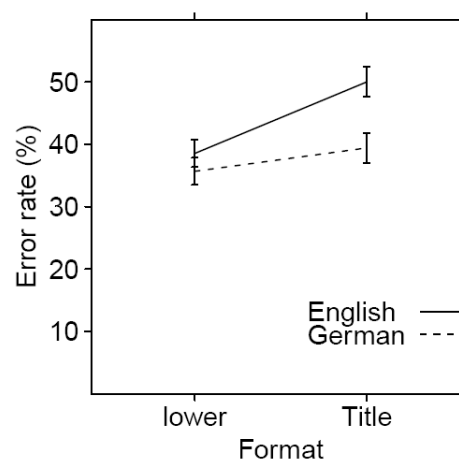
In the separate analysis of data from 200-level participants, Item Format interacts with Sentence Language (real words, $F(1,104) = 6.12, p < 0.01$). This interaction comes about because although error rates are not affected by Item Format after German contexts, presentation of a German word (noun) with an initial capital reduces the error rate after English contexts (see Graph 7). In contrast, when confronted with nonwords in Title (German-like) case, these participants are more likely to respond that the stimulus is a word ($F(1,155) = 86.99, p < 0.001$; see Graph 8).

Graph 5

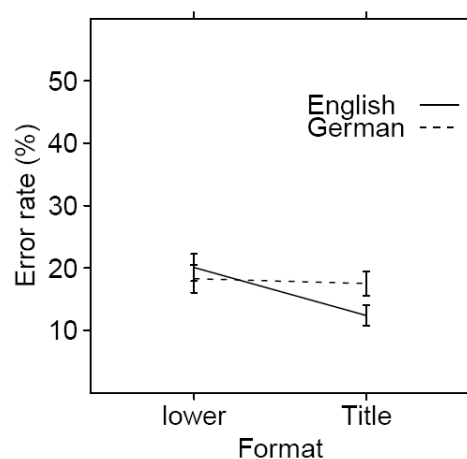
Mean incorrect responses to lower vs. Title case presentations of Interlingual Homographs and Real Words, for 100-level participants.

**Graph 6**

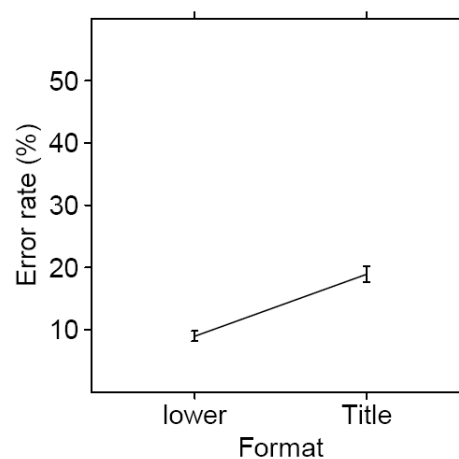
Mean incorrect responses to lower vs. Title case nonwords after English and German context sentences, for 100-level participants.

**Graph 7**

Mean incorrect responses after English and German context sentences, for lower vs. Title case targets, for 200-level participants.

**Graph 8**

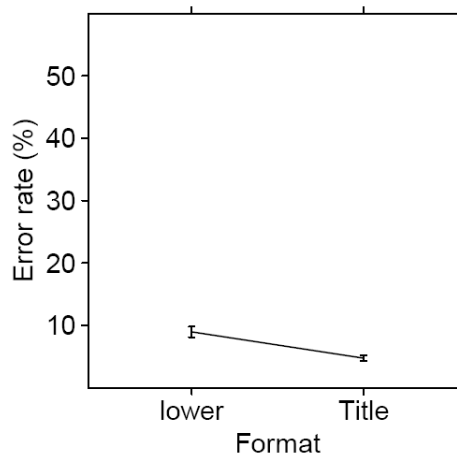
Mean incorrect responses to lower vs. Title case nonword targets, for 200-level participants.



Finally, data from 300-level participants show no main effects and no interactions of any factors in the analysis of real word errors; the only conventionally significant effect is for Item Format for errors to nonwords ($F(1,155) = 25.17, p < 0.001$; see Graph 9). Also noticeable is that error rates to the lower case items are similar to those observed for the 200-level participants above – what has changed is that there are now many fewer errors to the Title case items, even though the Item Format difference is still significant.

Graph 9

Mean incorrect responses to lower vs. Title case nonword targets, for 300-level participants.

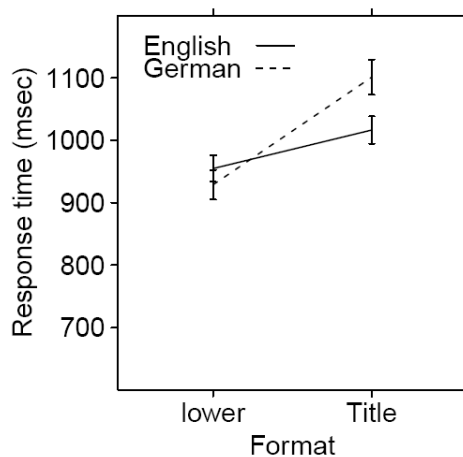


2.6 RT analysis

A general observation which can be made from looking at the response latencies (see Graphs 10-12), and which has been confirmed by statistical analyses, is that correct responses to real words are faster across all proficiency levels for lower case (for 100-level, $F(1,104) = 24.86, p < 0.001$; for 200-level, $F(1,104) = 45.19, p < 0.001$; for 300-level, $F(1,104) = 19.46, p < 0.001$).

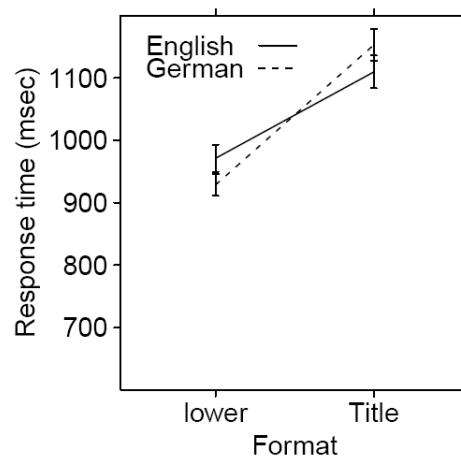
Graph 10

Mean response times to lower vs. Title case targets, after English and German context sentences, for 100-level participants



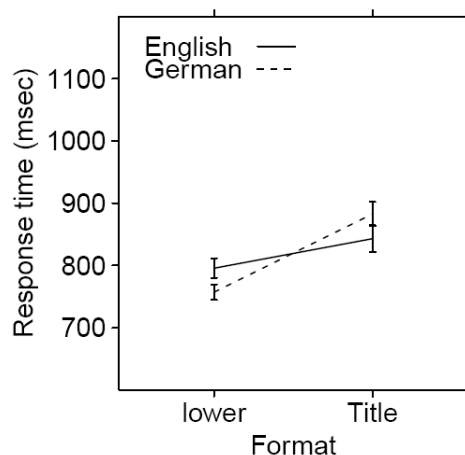
Graph 13

Mean response times to lower vs. Title case nonwords, after English and German context sentences, for 100-level participants

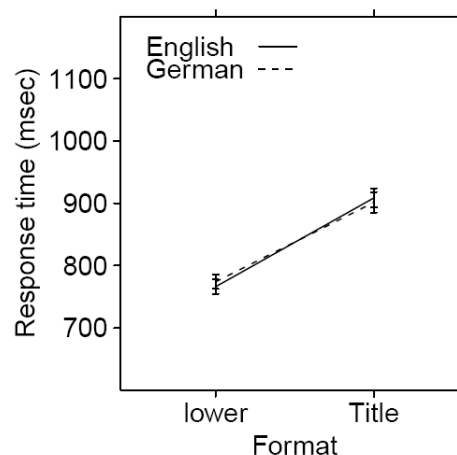


Graph 11

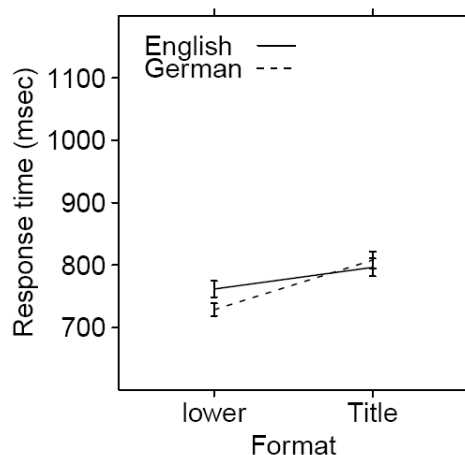
As above, for 200-level participants

**Graph 14**

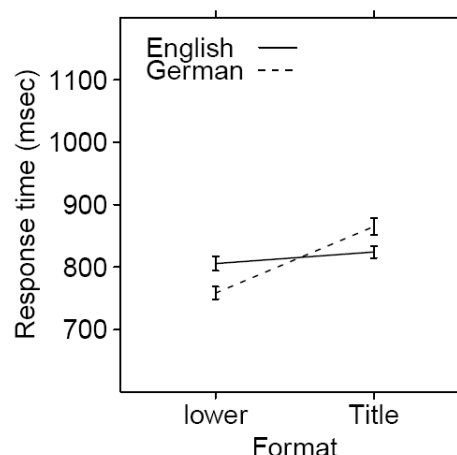
As above, for 200-level participants

**Graph 12**

As above, for 300-level participants

**Graph 15**

As above, for 300-level participants



Interestingly, the Item Format difference is consistently stronger in the German context across all levels (for 100-level, $F(1,104) = 10.39$, $p < 0.001$; for 200-level, $F(1,104) = 8.89$, $p < 0.003$; for 300-level, $F(1,104) = 4.70$, $p < 0.03$; Graphs 10-12). In addition, Item Type interacts with Item Format for low proficiency (100-level) participants ($F(1,104) = 6.46$, $p < 0.01$). That is, whereas response times to interlingual homographs are not affected by Item Format, presentation of a real German word with an onset capital reduces the speed with which the subjects can respond to it. A further observation is that the more proficient participants recognise interlingual homographs more rapidly than real German words (for 200-level, $F(1,104) = 6.81$, $p < 0.01$; for 300-level, $F(1,104) = 3.29$, $p = 0.07$).

As with real word responses, response times for correct rejections of nonwords show a significant Item Format effect across all proficiency levels (see Graphs 13-15), with Title case taking longer to reject (for 100-level, $F(1,155) = 63.83$, $p < 0.001$; for 200-level, $F(1,155) = 156.35$, $p < 0.001$; for 300-level, $F(1,155) = 52.74$, $p < 0.001$). At 100- and 300-level the Item Format effect interacted with Sentence Language (for 100-level, $F(1,155) = 5.85$, $p < 0.01$; for 300-level, $F(1,155) = 19.49$, $p < 0.001$), reflecting the

fact that it took these participants longer to reject nonsense words in Title case following German sentence primes (compare the two lines in Graphs 13 and 15).

3. Discussion

3.1 Error analysis

In line with expectations, the results presented above indicate that with increasing proficiency level, language-specific cues seem to have a stronger impact on visual word recognition processes, ie increasingly facilitating correct responses to real word items as well as correct rejections of nonwords. Another finding is that lower proficiency L2 speakers are more strongly influenced by their L1 vocabulary in making an L2 lexical decision response. This is shown in the data with participants with little exposure to German being more inclined to reject interlingual homographs as not being German words. This tendency becomes even stronger when these IHGs are presented with an onset capital letter. Interestingly, this problem does not arise with correct responses to matched controls (RWs), which also have the first letter capitalised. In addition, the same participants have more difficulties rejecting nonwords which have the first letter capitalised. This result is somewhat surprising, since our general expectation would be that onset capitalisation should facilitate the recognition of German words. One interpretation for the observed results is that low proficiency learners are not completely oblivious to (noun) capitalisation in German; they are simply being misled by the experimental requirements. Particularly with respect to nonsense words, this means that something unknown, but German-like in its spelling, is frequently reported as a German word. Another interpretation of the findings is that IHGs can be expected to remain stored as two separate entries with rather weak (if any) connections between the two languages. This interpretation would imply a developmental pattern of bilingual lexical organisation, with the two languages becoming increasingly separated.

In contrast, slightly more advanced bilinguals (our 200-level participants) show less L1 interference when responding to real words in the L2 (German). This is primarily reflected in the absence of a main effect distinguishing IHGs and matched RWs, and of any interactions involving this Item Type factor. A finding that is more in line with our expectations outlined above is that these participants' responses are affected by Item Format in a facilitatory manner. This is reflected in the interaction with contextual support. Whereas errors after German contexts are not affected by onset capitalisation, the particular format of stimulus presentation affects peoples' responses after English contexts; that is, it reduces the error rate. This result can only be explained by assuming that specific language cues can indeed set up a particular language mode, thereby facilitating real word responses. As was observed for low proficiency learners, slightly more advanced participants (200-level) are more prone to incorrectly accept as real words those German nonwords which are presented with the first letter capitalised. Again, this supports our view of language cues

having a strong impact on cognitive processing; in this case resulting in an inhibitory rather than facilitatory effect.

Finally, the error rate data for 300-level students do not reveal any main effects or interactions except for one: subjects make more errors when they have to reject a nonsense item with an onset capital. This is also observed with lower proficiency groups. Thus, the highly proficient participant group is clearly not being misled by the experimental factors to nearly the same extent as the less advanced learners; however what is evident is that all subjects are influenced by the fact that Title case marks nouns in German. Notably, the evident absence of a Sentence Language effect with increasing proficiency level is compatible with the idea that language-specific cues are processed bottom-up and largely independently of top-down cues from the context language or from the lexicon.

3.2 RT analysis

The response time analyses revealed a strikingly consistent response pattern across all three proficiency levels. This pattern occurs in both correct responses to real word items and correct rejections of nonwords. First, participants at all levels are slowed down when responding to German-biased items, ie those presented with an onset capital (although overall mean response times decrease significantly with increasing proficiency levels). This observation is confirmed statistically as a persistent effect of Item Format. A possible explanation of this response delay is that there is an additional consistency verification involved for an accessed German word, to ensure that the word is a noun (which requires capitalisation).² This conjecture seems to be supported by the second observation, namely that responses are more rapid when items appear after a German context and all in lower case. This is reflected in the statistical analysis as an interaction of Item Format and Sentence Language.

One final conventionally significant effect from the real word data is the Item Type effect. More specifically, IHGs are accepted more rapidly than their matched RW controls. Notably, this effect is also be found separately for 300- and 200-level participants, but not for 100-level, so it seems to be something that is connected to increasing proficiency. In line with previous research (Dijkstra, Timmermans et al 2000), this finding can be interpreted as a gradual cumulative effect of the bilinguals' two languages.

4. Conclusion

Embedding the findings presented above in the current research literature, we find that they not only further support common concepts and understanding of bilingual VWR, but also provide new insights into cognitive processes of a bilingual speaker.

² Please note that due to experimental restrictions not all real word items were nouns.

Addressing the question of the extent to which (sub-) lexical information can facilitate or inhibit bilingual VWR, recent research conducted by Vaid and French-Mestre (2002) suggests that bilinguals make use of certain language cues. The experimenters presented to their French-English subjects words that were either marked or unmarked for either L1 (French) or L2 (English) on the basis of digram frequency (eg OEUF for French, and KICK for English). The subjects' task was to decide which of these two languages the presented item belonged to. Participants' responses were faster for orthographically marked than unmarked words, particularly in the second language (English). The researchers interpreted these results in favour of a perceptual search strategy. That is, the recognition of orthographically marked words was facilitated because the late bilingual subjects (ie those who had learned English after the age of 12) employed bottom-up cues. We addressed this finding in the current research by investigating the role played by language-specific cues (in the form of onset capitals indicating German words) in English-German bilinguals' VWR. Our findings, based on error analyses, confirmed the previously observed facilitatory effects of language-specific cues on lexical access. However, we also extended the previous investigation by taking a related question into account, ie what effects might L2 proficiency have on the manifestation of facilitatory versus inhibitory dynamics? Interestingly, the effects reported above were more likely to be observed with more proficient bilinguals than their less bilingual counterparts. This finding is not surprising and does not contradict our expectations. Taken together with the evidence that language-specific cues are processed bottom-up and largely independently of the context language or the lexicon (as reflected in decreasing context language effects as proficiency level increases), the findings could also be interpreted in line with the BIA+ (Dijkstra and Van Heuven 2002) model. As explained above, the model assumes a gradual activation of sub-lexical, lexical, and conceptual levels during visual word recognition. The model also proposes a language tag/ node which can facilitate language selection. Let us assume that the sub-lexical level – and thus language-specific cues – is connected to a specific language tag or node. When confronted with an onset capital, a high proficiency learner might make quick use of a connection of this format to a particular language tag, informing him/ her about the language being processed and selected from; ie facilitating responses. A less proficient learner might not yet have established that connection, due to lower exposure to the L2 as well as a smaller vocabulary size. This would explain the facilitatory effect of language-specific cues being strongest for high proficiency learners.

It needs to be noted at this stage, however, that in contrast to our error rate results, response latencies across all levels revealed a delay for items marked 'German' (ie with an onset capital). We believe that this delay should not be read as an inhibitory effect. A more reasonable explanation of this delay is that there is an additional stage of consistency verification involved for an accessed German word, to ensure that the word is a noun, since only nouns require capitalisation. This conjecture is supported by the absence of any evident language context effect. That is, after German contexts, items presented in lower case were of all conditions the fastest responded to; however, items presented with an onset capital showed the reverse effect, ie being responded to the slowest of all conditions. This result does not allow for a clear-cut elaboration of the extent to which a bilingual can either use or discard sub-lexical

information in specific language contexts, or whether this information is sufficient to constrain (if not eliminate) lexical interaction between the bilingual's languages.

In line with this research direction, Libben and Titone (2009) have recently confirmed well-established claims made within monolingual word recognition research, arguing that bilingual lexical access at early stages of comprehension (ie bottom-up effects) is non-selective, but that selection from accessed words is rapidly resolved in semantically biased contexts at later stages of comprehension (ie top-down effects). Their claim was based on the lack of evidence of cognate facilitation or interlingual homograph interference for late-stage eye movement measures, but the opposite effect for early-stage comprehension measures. Considering our own results – on the one hand, error data suggest that sentence context is irrelevant to the processes involved in VWR, at least for more proficient speakers. On the other hand, response time data indicate a potential consistency verification process. This process would not and does not support fast responses in the most favourable and expected condition (ie to an item with an onset capital, embedded in a German context). Finally, the task has been performed in the participants' weaker L2 and it is possible that the specific language information is just not as readily available or of direct use to an L2 speaker.

Thus, not only is it important that future research directions address questions of the role of sub-lexical and/ or language-specific information, as well as proficiency level on visual lexical recognition processes, but also that more naturalistic experimental designs should evolve. Current research underway by the first author seeks to address some of the unresolved issues above and further confirm the recent findings by collecting data from German-English bilinguals.

Appendix: List of experimental items

Interlingual Homograph (IHG)	Control Filler (week one and week two)	
ALTER*	stein	karte
BAD	mut	arm
BALD*	wahr	dank
BITTEN*	stelle	nennen
BRIEF*	liebe	natur
CHEF*	knie	ewig
DOSE*	mehl	kamm
FASTEN*	ketten	rocken
GENIE	herab	busch
GIFT*	egal	sekt
GUT*	bis	was
HALL*	tote	nase
HANDY*	grund	punkt
HELL	hals	haut
HERD*	heim	heer
HERB*	kern	edel
HOSE*	tanz	topf
HUT*	los	lok
KIND*	dort	hoch
LIST*	faul	sofa
MADE	lamm	rahm
MIST*	ober	oase
MODE*	kauf	mord
MUSTER*	bitter	kochen
MUTTER*	fehlen	kosten
NOTE	sand	bier
NUN*	mai	uhr
RAT*	tor	rad
RATE	farm	wehr
ROMAN*	stoff	vogel
SAGE*	rost	maus
SMOKING*	frostig	dreckig
STERN*	fisch	kunde
STILL	braun	miete
STRAND*	nachts	teufel
TAG*	bau	all
TASTE*	beige	tanne
TELLER	kuchen	trauen
TOLL	zoll	matt

*Items used by Elston-Güttler and colleagues (2005).

References

- Andrews, S. 2006. Preface. In *From Inkmarks to Ideas: Current Issues in Lexical Processing*, S. Andrews (ed.). New York: Psychology Press, xix-xxix.
- Baayen, R., Piepenbrock, R. and Van Rijn, H. 1993. *The CELEX Lexical Database* [CD-Rom]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.
- Balota, D. A. 1994. Visual word recognition: The journey from features to meaning. In *Handbook of Psycholinguistics*, M. A. Gernsbacher (ed.). San Diego: Academic Press, 303-358.
- Brysbaert, M., Dyck, G. V. and Poel, M. V. D. 1999. Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception and Performance* 25(1), 137-148.
- Cambridge International Dictionary of English*, 1999. New York: Cambridge University Press.
- Chambers, C. G. and Cooke, H. 2009. Lexical competition during second-language listening: Sentence context, but not proficiency, constrains interference from the native lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35 (4), 1029-1040.
- Dijkstra, T. 2003. Lexical processing in bilinguals and multilinguals. In *The Multilingual Lexicon*, J. Cenoz, B. Hufeisen and U. Jessner (eds). Dordrecht: Kluwer, 11-26.
- Dijkstra, T. 2005. Bilingual visual word recognition and lexical access. In *Handbook of Bilingualism: Psycholinguistic Approaches*, J. F. Kroll and A. M. B. de Groot (eds). Oxford: Oxford University Press, 179-201.
- Dijkstra, T., Timmermans, M. and Schriefers, H. 2000. On being blinded by your other language: Effects of task demands on interlingual homograph recognition. *Journal of Memory and Language* 42(4), 445-464.
- Dijkstra, T. and van Heuven, W. J. B. 1998. The BIA model and bilingual word recognition. In *Localist Connectionist Approaches to Human Cognition*, J. Grainger and A. M. Jacobs (eds). Mahwah, NJ: Lawrence Erlbaum, 189-225.
- Dijkstra, T. and Van Heuven, W. J. B. 2002. The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition* 5(3), 175-197.
- Duyck, W., Assche, E. V., Drieghe, D. and Hartsuiker, R. J. 2007. Visual word recognition by bilinguals in a sentence context: Evidence for nonselective

- lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33(4), 663-679.
- Duyck, W., Desmet, T., Verbeke, L.P.C. and Brysbaert, M. 2004. Wordgen: A tool for word selection and nonword generation in Dutch, English, German, and French. *Behavior Research Methods, Instruments & Computers* 36, 488-499.
- Elston-Güttler, K. E., Gunter, T. C., and Kotz, S. A. 2005. Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research* 25(1), 57-70.
- Green, D. W. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1(2), 67-81.
- Grosjean, F. 2001. The bilingual's language modes. In *One Mind, Two Languages: Bilingual Language Processing*, J. L. Nicol (ed.). Malden, MA: Blackwell, 1-22.
- Kroll, J. F. and Dijkstra, T. 2002. The bilingual lexicon. In *The Oxford Handbook of Applied Linguistics*, R. B. Kaplan (ed.). Oxford: Oxford University Press, 301-321.
- Kroll, J. F. and Stewart, E. 1994. Category interference in translation and picture naming: Evidence for asymmetric connection between bilingual memory representations. *Journal of Memory and Language* 33(2), 149-174.
- Lemhöfer, K. and Dijkstra, T. 2004. Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition* 32(4), 533-550.
- Libben, M. R. and Titone, D. A. 2009. Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 35(2), 381-390.
- Nas, G. 1983. Visual word recognition in bilinguals: Evidence for a cooperation between visual and sound based codes during access to a common lexical store. *Journal of Verbal Learning & Verbal Behavior* 22(5), 526-534.
- Perlmann-Balme, M. and Kiefer, P. 2002. *Start Deutsch. Deutschprüfungen für Erwachsene. A1-A2. Prüfungsziele, Testbeschreibung*. München, Frankfurt: Goethe-Institut und WBT.
- Schneider, W., Eschman, A. and Zuccolotto, A. 2002. *E-Prime Reference Guide*. Pittsburgh: Psychology Software Tools Inc.
- Schneider, W., Eschman, A. and Zuccolotto, A. 2002. *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools Inc.

- Schwartz, A. I. and Kroll, J. F. 2006. Bilingual lexical activation in sentence context. *Journal of Memory and Language* 55(2), 197-212.
- Schwartz, A. I., Kroll, J. F. and Diaz, M. 2000. Reading Spanish words with English word bodies: Activation of spelling-to-sound correspondences across languages. Paper presented at the Second International Conference on the Mental Lexicon., Montreal, Quebec, October 2000.
- Vaid, J. and Frenck-Mestre, C. 2002. Do orthographic cues aid language recognition? A laterality study with French-English bilinguals. *Brain and Language* 82(1), 47-53.
- van Hell, J. G. and Dijkstra, T. 2002. Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review* 9(4), 780-789.
- von Studnitz, R. E. and Green, D. 2002. Interlingual homograph interference in German-English bilinguals: Its modulation and locus of control. *Bilingualism: Language and Cognition* 5(1), 1-23.