Parallel language activation during word processing in bilinguals: Evidence from word production in sentence context*

PETER A. STARREVELD
Department of Psychology, University of Amsterdam
ANNETTE M. B. DE GROOT
Department of Psychology, University of Amsterdam
BART M. M. ROSSMARK
Department of Psychology, University of Amsterdam
JANET G. VAN HELL
Department of Psychology, Pennsylvania State University & Behavioural Science Institute, Radboud University Nijmegen

In two picture-naming experiments we examined whether bilinguals co-activate the non-target language during word production in the target language. The pictures were presented out-of-context (Experiment 1) or in visually presented sentence contexts (Experiment 2). In both experiments different participant groups performed the task in Dutch, their native language (L1), or in English, their second language (L2). The pictures' names were Dutch–English cognates or non-cognates, the cognate effect serving as the marker of activation of the non-target language. In Experiment 2 we also examined the effect of sentence constraint. In both experiments a cognate effect occurred, but it was modulated by language and sentence constraint: The effect was larger in L2 than in L1, larger with low-constraint sentences than with high-constraint sentences, and disappeared in the high-constraint L1 condition. These results extend earlier bilingual word-recognition and out-of-context production studies, suggesting that also during word production in context, co-activation of the non-target language occurs.

The outcome of the substantial body of studies that examined bilingual word processing led some authors to conclude that bilingual language processing is fundamentally language-nonselective. This means that bilinguals cannot switch off the contextually inappropriate language, even when task performance takes place under circumstances that are as unilingual as possible. Most of the evidence supporting this position derives from studies that examined how bilinguals recognize words presented in isolation rather than ones embedded in a meaningful sentence or discourse context. Far fewer studies have examined the language-(non)selective nature of bilingual language processing by looking at the way bilinguals recognize words in context, the natural habitat of words. Also, far fewer studies have examined word production in isolation than word recognition in isolation. As far as we

know, the present study is among the first to investigate bilingual word production in sentence context, thus testing the scope of the language-nonselective language processing claim. Before detailing our study, we will first review the prior evidence.

Word recognition

Evidence that the recognition of words presented VISUALLY AND IN ISOLATION is language-nonselective comes from three lines of study. The first exploits the fact that one and the same written word may map onto two different meanings in a bilingual's two languages (e.g., the English word *coin* means "corner" in French). Many studies have shown that bilinguals take longer to recognize such "interlexical homographs" than to recognize matched non-homographic control words (e.g., De Groot, Delmaar & Lupker, 2000; Dijkstra, De Bruijn, Schriefers & Ten Brinke, 2000; Dijkstra, Grainger & Van Heuven, 1999; Dijkstra, Van Jaarsveld & Ten Brinke, 1998; Jared & Szucs, 2002; Von Studnitz & Green, 2002). The second group of studies examines whether word recognition is affected by so-called "interlexical

---

Address for correspondence:
Peter A. Starreveld, Department of Psychology, University of Amsterdam, Weesperstraat 4, 1018 XA Amsterdam, The Netherlands, or Annette M. B. de Groot, same address.
P.A.Starreveld@uva.nl, A.M.B.deGroot@uva.nl

neighbors", words in the non-target language that closely resemble (but are not exactly identical to) a target-language word but mean something else (e.g., Dutch *mand* "basket", is an interlexical neighbor of English *sand*). It has repeatedly been found that interlexical neighbors indeed affect word recognition (Bijeljac-Babic, Biardeau & Grainger, 1997; Grainger & Dijkstra, 1992; Jared & Kroll, 2001; Van Heuven, Dijkstra & Grainger, 1998). The third relevant type of studies compares responses to cognates – translation pairs that, in addition to sharing meaning, largely or completely share form between a bilingual's two languages – with responses to matched non-cognates. These studies typically find that cognates are processed faster and more accurately than non-cognates (e.g., Dijkstra et al., 1998; Dijkstra et al., 1999; Gollan, Forster & Frost, 1997; Schwartz, Kroll & Diaz, 2007; Van Hell & Dijkstra, 2002). In all three types of studies the critical effects are attributed to language-nonselective processing: During visual word recognition, elements in the non-target language system are co-activated with elements in the target-language system.

Analogous evidence for language-nonselective recognition of SPOKEN words is obtained from studies employing the "visual world paradigm" (Blumenfeld & Marian, 2007; Ju & Luce, 2004; Marian & Spivey, 2003a, b; Spivey & Marian, 1999; Weber & Cutler, 2004). In this paradigm the participants are presented with a visual display placed in front of them (or on a computer screen), which contains a number of (pictures of) objects. The display might, for instance, contain a toy shark, a balloon, a toy horse, and a napkin, and a Russian–English bilingual participant may be orally instructed to "pick up the shark" (Marian & Spivey, 2003b). Eye-movement-recording apparatus registers where the participant looks at while carrying out this instruction. The crucial finding in these studies is that a non-target object (e.g., the balloon) whose name in the non-target language resembles the target object's name in the target language (*balloon* is *sharik* in Russian, resembling *shark*) is looked at significantly more often than non-target objects whose names in the non-target language do not resemble the target object's name. This finding suggests that a spoken word activates similarly sounding words in the non-target language. However, while some of these studies (Marian & Spivey, 2003a; Spivey & Marian, 1999) obtained this effect both when relatively strong L1 (the native language) and weaker L2 (the second language) served as the target language, other studies (Blumenfeld & Marian, 2007; Marian & Spivey, 2003b; Weber & Cutler, 2004) only obtained the effect with weaker L2 as the target language, indicating that relative proficiency in the two languages can modulate the influence of the other language on word recognition.

The few studies in which the words to be recognized were presented visually and IN CONTEXT show a mixed pattern of results. The in-context studies in which the above cognate and/or homograph manipulations were used tend to show that the homograph effects generally disappear (Elston-Güttler, Gunter & Kotz, 2005; Schwartz & Arêas Da Luz Fontes, 2008; Schwartz & Kroll, 2006; Titone, Libben, Mercier, Whitford & Pivneva, 2011), even when weaker L2 is the target language (but see Libben & Titone, 2009), suggesting that sentence context can constrain activation to the contextually-appropriate language. In contrast, cognate effects still occur when the critical words are presented in a sentence context (Duyck, Van Assche, Drieghe & Hartsuiker, 2007; Libben & Titone, 2009; Schwartz & Kroll, 2006; Titone et al., 2011; Van Assche, Drieghe, Duyck, Welvaert & Hartsuiker, 2011; Van Assche, Duyck, Hartsuiker & Diependaele, 2009; Van Hell & De Groot, 2008), especially when the contexts are "low-constraint", meaning that the sentence context preceding the target word does not severely constrain the set of possible target words (e.g., *She gave the guest a spoon so he could eat his pudding*, with *spoon* being the target word). In contrast, when the sentence contexts are "high-constraint", strongly constraining the set of possible target words (e.g., *One usually eats soup with a spoon because it is a lot trickier with a fork*, with again *spoon* as the target word) often no cognate effect is obtained (but see Van Assche et al., 2011). The cognate effect obtained with low-constraint sentences may even appear when strong L1 is the test language (Titone et al., 2011; Van Assche et al., 2009). Finally, Chambers and Cooke (2009) showed a role of sentence constraint in L2 SPOKEN-word recognition, using a sentence–context version of the visual world paradigm presented earlier.

The above discussion warrants the conclusion that bilingual word recognition is often language-nonselective but also suggests that relative language proficiency, stimulus type and linguistic context type qualify language-nonselective word recognition. Support for language-nonselective language processing as obtained from word-recognition studies does, however, not necessarily generalize to word production.

Word production in isolation

Word production differs in one, potentially crucial, aspect from word recognition: While the reader/listener is driven by externally presented information, the speaker intentionally chooses the target language and may perhaps exert control on which representations become activated (Costa & Santesteban, 2004). In fact, because the reader/listener is driven by externally presented information, in one sense the evidence of language-nonselective bilingual language processing as derived from word-recognition studies is rather trivial: Word recognition in fluent language users is known to be a highly automated process. Therefore, if a bilingual
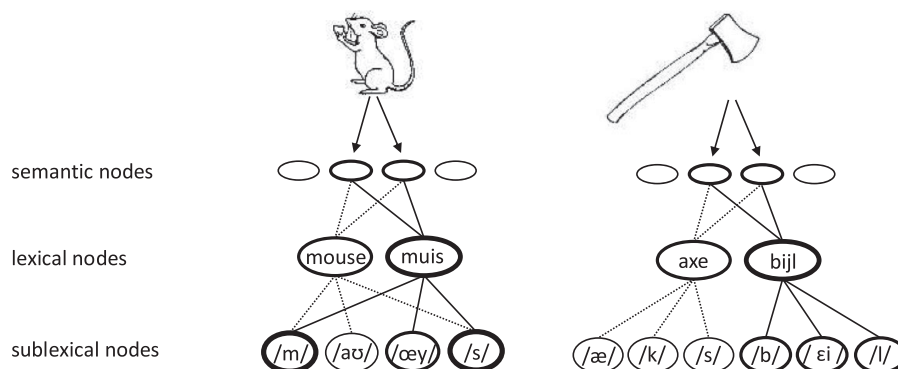
Figure 1. A model of picture naming in bilinguals. All connections shown in the model are bidirectional, thicker ellipses represent more activation. The left and right panels illustrate the activation of the memory representations of the (English–Dutch) cognate pair *mouse–muis* and the non-cognate pair *axe–bijl* when the picture is named in Dutch and activation has just reached the sublexical nodes. Subsequently, the lexical node of the cognate *muis* will receive more feedback from the sublexical layer to the lexical layer than the lexical node of the non-cognate *bijl*. As a result, with the continued spread of activation through the network, the complete set of sublexical nodes representing a cognate will receive more activation than the one representing a non-cognate. Based on Costa et al. (2000).

is fluent in the two languages he or she masters, an externally presented word stimulus will automatically trigger memory units in both language subsystems. In contrast, word-production studies can use non-verbal stimuli that do not automatically trigger memory units related to language. For this reason, to inform and resolve the present discussion, evidence from bilingual word-production is indispensable.

All word-production studies that have hitherto addressed the present question about the (non)selectivity of bilingual word processing have examined word production in isolation. The participants always performed some version of the picture-naming task. In a subset of these studies the pictures' names in the participants' two languages were either cognates or non-cognates (e.g., Christoffels, De Groot & Kroll, 2006; Costa, Caramazza & Sebastián-Gallés, 2000; Hoshino & Kroll, 2007). Despite the fact that the written or spoken forms of the critical words, which define cognate status, were not actually present during stimulus presentation, pictures with cognate names were named faster than those with non-cognate names. The effect occurred both with picture naming in the stronger L1 and in the weaker L2, although it appeared to be larger in L2. The favored interpretation of this cognate effect (see Figure 1, based on Costa et al., 2000) is that, starting with the initial activation of a semantic representation that is shared between the two languages, bilingual word production is language-nonselective all the way down to a layer of sublexical phonological units ("nodes") in the word-production system, this layer being shared between a bilinguals' two languages. As a consequence, the set of sublexical phonological nodes that represent cognate names accumulates activation from two previously

activated lexical nodes, one for each language (Figure 1, left) and naming can proceed relatively rapidly (as compared with naming non-cognates, whose sublexical phonological representations only receive activation from one lexical node; see Figure 1, right).

In another version of the picture-naming task not only the picture to name but also a distracter stimulus is presented on a trial and typically only non-cognate stimulus materials are used. In most of these studies (Costa, Colomé, Gómez & Sebastián-Gallés, 2003; Costa, Miozzo & Caramazza, 1999; Hermans, Bongaerts, De Bot & Schreuder, 1998), the distracters were (spoken or written) words in the target or non-target language that shared some specific relation (e.g., semantic or phonological) with the picture's name in the target language (e.g., a picture of a mountain, to be named in English, being paired with the distracter word *bench*, whose sound resembles *berg*, Dutch for "mountain"; Hermans et al., 1998). In Colomé and Miozzo (2010), both the targets and the distracters were pictures. These studies have produced evidence of language-nonselective word production by showing differences in picture-naming latencies between the experimental conditions (e.g., a picture of a mountain with "bench" as distracter) and a control condition (a picture of a mountain with an unrelated word as distracter).

Finally, in two bilingual picture-naming studies, the presented pictures, all with non-cognate names, had to be named covertly (instead of overtly) en route to the response (Colomé, 2001; Rodriguez-Fornells, Van der Lugt, Rotte, Britti, Heinze & Münte, 2005). In Colomé's study, which tested Catalan–Spanish bilinguals, each picture was accompanied by a phoneme (represented by its corresponding letter). The participants' task was to

indicate whether or not the phoneme was part of the Catalan name of the picture (this procedure is called "phoneme monitoring"). The critical finding was that "no" responses were slower when the phoneme presented on that trial was part of the picture's name in Spanish (but not in Catalan; hence the "no" response) than when it occurred neither in the picture's Catalan name nor in its Spanish name. This finding suggests that the picture's name is phonologically encoded in both languages and, thus, that out-of-context bilingual word production is language-nonselective. The activated Spanish name, which contains the phoneme presented on that trial, causes interference and delays the response (but see Hermans, Ormel, Van Besselaar & Van Hell, 2011). Using a modified version of this paradigm, Rodriguez-Fornells et al. (2005) obtained a similar result.

The present study: Word production in sentence context

The evidence presented so far suggests that language-nonselective bilingual word processing holds for word recognition, both in isolation and in context (at least for the in-context studies that exploited the cognate manipulation), and for word production in isolation. In addition, it suggests that language nonselectiveness is modulated by the bilingual's relative proficiency in the two languages and by sentence constraint. In order to complete the picture and be able to make a more general claim regarding the language-(non)selective nature of bilingual word processing, word production in sentence context needs to be studied. This is what we have done here, using the common picture-naming task and comparing performance to pictures with cognate names and non-cognate names.

We used the cognate manipulation as marker of language-(non)selective processing despite the fact that its suitability is under discussion. Several authors, including some of us, have suggested that cognates and non-cognates may be represented in bilingual memory in a qualitatively different way: A pair of cognates may share their meaning representations while a pair of non-cognates may not do so (De Groot & Nas, 1991), or the former may share larger parts of their meaning representations than the latter (Van Hell & De Groot, 1998); cognates may share a morphological representation in memory while non-cognates do not (Sánchez-Casas, Davis & García-Albea, 1992); and cognates may share a word-form representation in memory while non-cognates do not (Dijkstra & Van Heuven, 2002). If true, cognate effects may arise because cognates have a higher functional frequency than non-cognates (see Titone et al., 2011, for a discussion). In addition, the inclusion of pictures with cognate names *per se* may induce parallel activation in both language systems (Hermans et al., 2011), perhaps because the participants

notice the presence of cognates and consequently move away from a monolingual towards a bilingual mode of processing, co-activating the non-target language (e.g., Grosjean, 2001).

Despite these potential problems with using the cognate manipulation we knowingly opted for it for three reasons. First, it appears to be the most frequently exploited manipulation in studies on the language-(non)selective nature of bilingual word processing and this certainly holds for the in-context studies. Using the same stimulus manipulation would enable us to compare the effects obtained in these previous studies with those obtained here. Second, the other stimulus manipulations used in the word-recognition studies reviewed above, those that make use of interlexical homographs or interlexical neighbors, are unsuitable for word-production studies because, unlike a pair of cognates, interlexical homographs and interlexical neighbors do not share meaning between a bilingual's two languages. When presenting a picture of, say, a coin to an English–French bilingual, the production process starts off with the activation of the meaning representation of a coin, but NOT with the meaning representation of French *coin* (recall that *coin* means "corner" in French). The sublexical phonological representation of the picture's English name *coin* could thus never receive converging activation from the meaning of the French word *coin*. Third, the phoneme-monitoring task that has previously been used in bilingual word-production studies (Colomé, 2001; Hermans et al., 2011; Rodriguez-Fornells et al., 2005; see above) and which may use pictures with non-cognate names exclusively is clearly a less natural task than simple picture naming.

As far as we know, this is the first study that examines bilingual word production in sentence context using picture naming, although a study by Van Hell and De Groot (2008) can be argued to be quite similar to the present one. In that study, two experiments were included in which visually presented target words had to be translated into the other language. The cognate status of the target words was manipulated, and the target words were either presented in isolation or in a sentence context that was either high-constraint or low-constraint. In both translation directions (from L1 to L2 and vice versa) a cognate effect occurred, but the effect in the high-constraint context condition was substantially smaller than in the low-constraint context condition and in the isolated word condition. If word translation is a form of word production, as some have argued (De Groot, 2011, pp. 259–265; La Heij, De Bruyn, Elens, Hartsuiker, Helaha & Van Schelven, 1990; Miller & Kroll, 2002), these findings suggest that bilingual word production in context is language-nonselective. However, two alternative interpretations of these results can be offered. First, in word-translation studies the observed

cognate effects may have been caused by processes related to the recognition of the stimulus word rather than to processes related to the production of the target. Second, by its very nature, word translation plausibly always requires the parallel activation of the bilingual's two language subsystems. It is possible that the cognate effects were somehow caused by this aspect of the task.

The present study encompassed two experiments in which the same sets of pictures were used. The pictures' names were either cognates (25%) or non-cognates (75%). In Experiment 1 the pictures were presented in isolation whereas in Experiment 2 they were presented in a sentence context. The purpose of Experiment 1 was to replicate the common cognate effect in picture naming in isolation and to thus serve as a baseline for Experiment 2. The participants in both experiments were Dutch–English bilinguals with L1 Dutch stronger than L2 English. To be able to examine the role of language proficiency on language-(non)selectivity, in both experiments different groups of participants named the pictures in either L1 Dutch or in L2 English.

In Experiment 2 we presented rebus sentences (Potter, Kroll, Yachzel, Carpenter & Sherman, 1986) in which a noun was replaced by a picture. The sentences were presented visually, word by word, and the participants determined the presentation speed themselves (that is, we used a "self-paced" reading technique). Participants read the words silently and named the picture when it appeared. This task thus involves a switch from sentence comprehension to word production, a procedure used before in several studies on monolingual word production (e.g., Bock & Miller, 1991; Brown, 1979; La Heij, Starreveld & Steehouwer, 1993; Starreveld & La Heij, 1995). To examine a possible role of sentence constraint, both high-constraint and low-constraint sentence contexts were included. The procedure and the design of Experiment 2 were thus very similar to those used in the in-context bilingual word-recognition studies discussed above, especially those that exploited the cognate manipulation, the crucial difference being that we examined the effect of visual sentence context on word production.[1]

---

[1] Experiment 2 is not intended to be an experimental analogue of the most typical form of normal speech production, which is when a speaker addresses his or her interlocutor(s) with speech fragments larger than single words. Such speech starts off with a SPEAKER-INTERNAL process of conceptualizing the message to express in words. Developing experimental analogues of such speech is notoriously hard because it is impossible to gain experimental control over the myriad of possible utterances. In the present paradigm the participants also conceptualize a message prior to starting word production, but this process of conceptualization has no speaker-internal source but results from processing external linguistic and pictorial input. Our paradigm does, however, resemble one specific form of natural speech production, namely when in turn-taking in a conversation the person

taking over suggests a single candidate word to a conversation partner who pauses to search for a word. In both that situation and in our paradigm a single word is conceived of and spoken after a prior process of conceptualization taking place on the basis of external language input.

Experiment 1: Picture naming in isolation

Method

Participants
Forty six students from the University of Amsterdam participated, 23 in Condition Dutch and 23 in Condition English. They responded to an advertisement posted within the university that asked students who had Dutch as their mother tongue to sign up. Members from this population generally share the following characteristics: They started preparatory training in English at primary school around age 10 years. At secondary school, from around age 12 years, they received between three and four hours of English training per week until their university enrollment around age 18. Their current university training requires them to read mainly English textbooks and articles and they are exposed to English daily through popular media. Previous studies in which participants from this same population were used have shown that their English is generally well developed but not as strongly as Dutch, as for instance is evident from slower responses in English than in Dutch on both word comprehension tasks (e.g., De Groot, Borgwaldt, Bos & Van den Eijnden, 2002) and word production tasks (e.g., Christoffels et al., 2006). As we shall see, the results of the present experiment also show this familiar pattern.

Materials and apparatus
One hundred pictures were selected from the corpus of black-and-white line drawings made available by Székely, D'Amico, Devescovi, Federmeier, Herron, Iyer, Jacobsen and Bates (2003). Fifty of these served as critical pictures. The names of 25 of these were Dutch–English cognates (seven of them identical, the remaining 18 non-identical). These pictures served as materials in the cognate condition. The names of the remaining 25 critical pictures were Dutch–English non-cognates. These pictures were used in the non-cognate condition. The remaining 50 pictures, all with non-cognate names, served as fillers and were also presented in warm-up trials and following an error (see Procedure below for details). Categorization of the critical pictures' names as cognates or non-cognates was first done by the experimenters on the basis of a subjective assessment of the orthographic and phonological similarity between the pictures' Dutch and English names. Subsequently, the critical items' cognate status was determined objectively

Table 1. *Means of characteristics of the pictures'
cognate and non-cognate names in both language
conditions (critical stimuli only; standard deviations in
parentheses).*

Variable	Cognates	Non-cognates
Dutch		
Number of syllables	1.16 (0.37)	1.04 (0.20)
Number of phonemes	3.68 (0.56)	3.72 (0.54)
Log frequency	1.72 (0.58)	1.64 (0.58)
Subjective frequency	4.38 (1.32)	4.45 (1.14)
English		
Number of syllables	1.20 (0.50)	1.36 (0.57)
Number of phonemes	3.76 (0.72)	4.00 (1.22)
Log Frequency	1.73 (0.63)	1.65 (0.55)
Subjective frequency	4.73 (1.03)	4.69 (0.92)

by applying Van Orden's (1987) formulae for orthographic and graphemic similarity to the translation pairs (see Appendix A). The average orthographic similarity scores for cognate and non-cognate pairs were .70 ($SD = .24$) and .11 ($SD = .10$), respectively (these scores vary between 0 and 1). The average graphemic similarity scores for cognate and non-cognate pairs were 697 ($SD = 236$) and 107 ($SD = 102$), respectively. Independent samples *t*-tests performed on the similarity scores showed both differences between the cognate pairs and the non-cognate pairs to be significant (orthographic similarity: $t(48) = 11.48$, $p < .001$; graphemic similarity: $t(48) = 11.46$, $p < .001$).

The critical pictures' names in the cognate and non-cognate conditions were matched on log word frequency (per million words in the CELEX databases, which are based on corpora of written texts; http://celex.mpi.nl/), number of syllables, and number of phonemes. In addition, for the cognate and non-cognate conditions, log word frequency, number of syllables, and number of phonemes were matched between languages. A series of *t*-tests confirmed that the matching procedure was successful (all ts > .10). This held for each of the languages separately and across the two languages. Table 1 shows the corresponding means and standard deviations. A complete list of the critical pictures' names in Dutch and English is presented in Appendix B. Given the constraints on stimulus selection, no other pictures with cognate names could be selected from the available set of pictures.

After the data of Experiments 1 and 2 had been collected, we had two separate groups of participants (from the same population as those tested in Experiments 1 and 2) give subjective frequency-of-use ratings to the names of all 100 pictures.[2] One group (N = 26) rated the pictures' Dutch names; the second (N = 28) rated their English names. All 100 words were presented mixed in a booklet consisting of 10 pages, 10 words per page, the page order differing between participants. Next to each word was a seven-point scale on which the participants indicated how often they thought they used the word (1 = seldom; 7 = very often). For each of the 50 critical picture names and for both languages an average score was subsequently calculated and these scores were used in an analysis of variance (ANOVA) with language (two levels: Dutch and English) and cognate status (two levels: cognates and non-cognates) as between-items variables. Neither the main effect of language ($p = .19$), nor the main effect of cognate status ($p = .94$), nor the interaction between these variables ($p = .82$) were significant. In other words, the four sets of picture names were not only matched on number of syllables, number of phonemes, and log word frequency but also on subjective frequency. The corresponding means and standard deviations are shown in Table 1.

The experiment was performed using Presentation® software (Version 9.90, www.neurobs.com). Pictures were presented on a fast cathode ray tube monitor running on a refresh rate of 70 Hz. Responses were collected using a voice key and were measured to the nearest millisecond.

Procedure
The participants were tested in a quiet, dimly lit room and were seated in front of a computer monitor at a comfortable viewing distance. All communication with the participants was in Dutch in Condition Dutch and in English in Condition English. The participants first read an instruction sheet after which they signed an informed consent form (the text was in the language of that condition). Next, in agreement with a common practice in picture-naming experiments (e.g., Costa et al., 1999; Glaser & Düngelhoff, 1984; Starreveld & La Heij, 1995, 1996), to familiarize the participants with the experimental materials prior to actual testing, the participants studied a booklet that contained all pictures (both critical and filler) and their names, the latter being printed below the pictures in the language of that condition. A familiarization phase has been shown to produce a more homogenous data set in terms of less variability in mean RTs and fewer errors in a following experimental phase, without affecting the basic characteristics of picture-naming processes (Alario, Ferrand, Laganaro, New, Frauenfelder & Segui, 2004). Critical cognate pictures, critical non-cognate pictures, and filler pictures were presented in a random order.

---

[2] This rating study was suggested to us by a reviewer of an earlier version of this text, which is why we only conducted it after data collection rather than before.

Participants were asked to study the booklet carefully and to use the provided names in the upcoming naming part of the experiment (instead of using any alternative word to name the picture).

After studying the booklet, the picture-naming part of the study started. The participants were asked to name the pictures (in the appropriate language) as quickly and accurately as possible. This picture-naming part started with the presentation of 20 warm-up trials that allowed participants to get used to the task and the experimental setting. Warm-up trials contained pictures that were randomly selected from the 50 filler pictures. Subsequently, all 100 selected pictures (critical and filler) were presented, the presentation sequence being randomized anew for each participant. The pictures were presented in four groups with a short break between the groups.

Each picture-naming trial involved the following sequence. A fixation point (a "+" sign) was presented in the centre of the screen for 500 ms. Then a blank screen was presented for 500 ms, followed by the presentation of the picture. The picture remained on the screen until the participant responded or until 2500 ms had elapsed. The experimenter then typed in a code indicating whether the response was correct or false or whether the voice key had malfunctioned. When an error was made, the experimenter provided the correct name to the participant. Finally, a blank screen was presented for 500 ms after which the next trial began. The time duration between picture onset and the moment the participant initiated the vocal response was registered as response time.

To reduce variance in the data, an additional trial was presented after trials in which participants made an error or the voice key malfunctioned. Such additional trials used pictures that were randomly selected from the filler pictures. Data obtained for warm-up trials and additional trials were not recorded.[3]

Results

Data collected for the filler pictures were not analyzed. For the remaining data, first, RTs from incorrect responses and from trials in which the voice key malfunctioned were removed. Next, RTs faster than 300 ms and slower than 2000 ms were removed. Finally, for each participant, RTs that deviated more than 3 SDs from the mean of each condition were removed. For Condition Dutch these exclusions accounted for 3.9%, 3.7%, 0.3%, and 1.6%

Table 2. *Participant mean reaction times (in milliseconds) per condition and error percentages (in parentheses) for Experiment 1.*

Condition	Cognates	Non-cognates	Cognate effect
Dutch	719 (4.2)	754 (3.7)	35°;♦
English	848 (7.8)	999 (10.3)	151°;◇
Language effect	129	245	

°$p < .01$ in the subject analysis; ♦$p < .05$ in the item analysis; ◇$p < .01$ in the item analysis

Note: Cognate effect = Non-cognates RT – Cognates RT. Language effect = English RT – Dutch RT. The significance levels for the cognate effects were calculated using one-sided *t*-tests.

of the data, respectively, while for Condition English they accounted for 9.0%, 15.4%, 1.1%, and 1.0% of the data, respectively.[4] The remaining RTs were used for the calculation of the means for the various conditions (see Table 2).

An ANOVA was performed on the mean RTs per participant per condition, with language (two levels: Dutch and English) as between-subjects variable and cognate status (two levels: cognates and non-cognates) as within-subjects variable. The corresponding item analysis was also performed, with language (two levels: Dutch and English) and cognate status (two levels: cognates and non-cognates) as between-items variables. Note that because cognate status is a within-subjects variable in the by-subjects analysis, the means that enter the analysis are averaged over pictures and the nuisance variance caused by differences in response speeds for different pictures within a condition is excluded. However, because cognate status is a between-items variable, the means that enter this analysis are averaged over participants and the nuisance variance caused by differences in response speeds for different pictures within a condition cannot be excluded. Therefore, the power to detect a cognate effect is much higher in the subject analysis than in the item analysis. The only way to increase the power to detect a cognate effect in the item analysis would be to increase the number of pictures in the cognate and non-cognate conditions, but this was impossible given the constraints on stimulus selection: As mentioned before, no further pictures from the initial set (Székely et al., 2003) met all selection criteria described earlier.

---

[3] After presenting all 100 selected pictures for naming once, three further presentation series followed in each of which all 100 pictures were again presented for naming. This was done to examine whether task practice might modulate the effect of cognate status and language. The results of this manipulation will be summarized in footnote 5 below.

[4] Name-retrieval difficulties may cause both wrong responses (resulting in common errors) and softly spoken correct responses (resulting in voice-key errors). Our participants probably experienced more name-retrieval difficulties when naming in English, their nonnative language, than in native Dutch. This can explain why we obtained both more common errors and more voice-key errors in Condition English than in Condition Dutch.

The effect of language was significant, $F_1(1,44) = 25.68$, $MSE = 31,343$, $p < .001$, and $F_2(1,96) = 104.39$, $MSE = 7,202$, $p < .001$, with English responses taking longer than Dutch responses. The effect of cognate status was also significant, $F_1(1,44) = 86.72$, $MSE = 2,285$, $p < .001$, and $F_2(1,96) = 26.43$, $MSE = 7,202$, $p < .001$, pictures with cognate names being named faster than pictures with non-cognate names. Finally, the interaction between language and cognate status was significant, $F_1(1,44) = 33.42$, $MSE = 2,285$, $p < .001$, and $F_2(1,96) = 9.16$, $MSE = 7,202$, $p = .003$, the cognate effect being larger when the pictures were named in English than when they were named in Dutch.[5] To evaluate the significance of the simple effects of cognate status, the obtained interaction was further examined by using planned comparisons that consisted of one-sided paired (in the subject analysis) and unpaired (in the item analysis) *t*-tests. The results of these analyses are listed in Table 2 above. Both languages showed a cognate effect.

To see whether speed–accuracy trade-offs may have occurred, the corresponding error analyses were performed on the number of errors per condition (see Table 2 for the corresponding error percentages). The effect of language was significant, $F_1(1,44) = 17.28$, $MSE = 2.19$, $p < .001$, and $F_2(1,96) = 8.15$, $MSE = 4.27$, $p = .005$. More errors were made in Condition English (9.1%) than in Condition Dutch (3.9%). The effect of cognate status was not significant, both $Fs < 1$, indicating that the same number of errors was made for pictures with cognate names as for pictures with non-cognate names. Finally, the interaction between language and cognate status was not significant, $F_1(1,44) = 2.06$, $MSE = 1.52$, $p = .16$, and $F_2 < 1$. We can conclude that no speed–accuracy trade-offs were apparent in the data.

Discussion

We found a clear effect of cognate status, both when Dutch–English bilingual participants named pictures in L1 Dutch and when they named them in L2 English. In both cases participants named pictures with cognate names faster than pictures with non-cognate names. The effect was considerably larger when the pictures were

named in L2 English than when they were named in L1 Dutch, but also in L1 Dutch the effect was significant.

Experiment 2: Picture naming in context

In this experiment we examined whether the cognate effect would also surface when participants named the pictures in sentence context.

Method

Participants

Forty-six students from the University of Amsterdam participated, 23 in Condition Dutch and 23 in Condition English. All participants were drawn from the same population as used in Experiment 1 but none of them had participated in that experiment. In an exit questionnaire the participants indicated their reading and speaking abilities in Dutch and English. They were asked to use a scale of 1–7, where a score of 7 indicated that they estimated their level of proficiency to be excellent while a score of 1 indicated they regarded it to be very poor. The scores confirmed that the bilinguals were more fluent in Dutch than in English. The average reading ability scores of the participants in Condition Dutch were 6.43 and 6.00 for Dutch and English, respectively. The analogous scores provided for Condition English were 6.18 and 5.62. The average speaking ability scores of the participants in Condition Dutch were 6.65 and 5.78 for Dutch and English, respectively. The analogous scores provided for Condition English were 6.49 and 5.50. A set of *t*-tests confirmed that all four of these contrasts were significant ($p < .05$), thus indicating that our participants' bilingualism was unbalanced for both reading and speaking. A further set of *t*-tests confirmed that the participants in Condition Dutch and Condition English did not differ from one another on any of the four measures (reading and speaking in English and Dutch), showing that the two samples of participants tested in the two language conditions were comparable.[6]

Materials

The same 50 critical and 50 filler pictures were used as in Experiment 1. However, the pictures were now embedded in visually presented sentences. Each critical picture was to be presented twice in each language condition, once in a high-constraint condition and once in a low-constraint condition. We therefore constructed a total of 200 critical sentences, 100 per language

---

[5] As mentioned in Footnote 3, to examine a possible influence of task practice, three further naming series of all 100 pictures followed the first. The data showed that participants' responses got faster over consecutive series. This effect was significant, $F_1(3,132) = 151.36$, $MSE = 2,764$, $p < .001$, and $F_2(3,288) = 292.85$, $MSE = 1,457$, $p < .001$. The interaction between language and series was also significant, $F_1(3,132) = 13.95$, $MSE = 2,764$, $p < .001$ and $F_2(3,288) = 21.46$, $MSE = 1,457$, $p < .001$, showing that the practice effect was larger in Condition English than in Condition Dutch. Finally, the interaction between cognate status and series was significant, $F_1(3,132) = 11.16$, $MSE = 1,295$, $p < .001$ and $F_2(3,288) = 8.94$, $MSE = 1,457$, $p < .001$, the cognate effect becoming smaller over consecutive series.

[6] The participants in Experiment 1 had not been asked to rate their reading and speaking abilities in Dutch and English after data collection but the data of Experiment 1 – shorter response times and fewer errors in Dutch than in English – clearly confirm our assertion that their L1 Dutch was stronger than their L2 English.

Table 3. *Examples of critical sentences used in Condition English in Experiment 2. The word in parentheses is presented in the form of the picture to be named.*

Sentence type	Stimulus type	Example sentence
High-constraint	Cognates	Popeye the sailorman has a tattoo of an (ANCHOR) on his arm.
		She went to the library to borrow a (BOOK) so she could study it at home.
	Non-cognates	The alpinist climbed the highest (MOUNTAIN) because he loved challenges.
		The goods were delivered in a cardboard (BOX) which I could barely carry upstairs.
Low-constraint	Cognates	In the middle of the square was an (ANCHOR) with a thick chain attached to it.
		I was deeply moved by the (BOOK) in which she described her youth.
	Non-cognates	Along the route we have to go around the (MOUNTAIN) because we cannot cross it.
		On the little black table was a (BOX) with the logo of the company on it.

condition. Of the 100 sentences per language condition, 50 constituted the high-constraint condition and the remaining 50 constituted the low-constraint condition. Within each of these groups of 50 sentences, 25 embedded a cognate and the remaining 25 embedded a non-cognate. In the high-constraint condition the pictures' names (and the associated concepts) were highly predictable from the prior sentence context, whereas in the low-constraint condition they were not (see the results of a norming study, below). For example, in the high-constraint condition a picture of spoon was embedded at the position of the underscores in the sentence: "One usually eats soup with a ___ because it is a lot trickier with a fork". In the low-constraint condition, the same picture of a spoon was embedded at the position of the underscores in the sentence: "She gave the guest a ___ so he could eat his pudding".

To guarantee that any effects of the critical variables (language, cognate status, sentence constraint) would not be contaminated by variability in sentence-processing difficulty due to differences in sentence structure, all critical sentences had similar grammatical structures: The first part of the sentence was always a main clause in which the final word was replaced by the picture to be named. The sentence then continued with a subordinate clause, another main clause, or a prepositional phrase. The picture to be named always appeared around the middle of the complete sentence. In addition, the length of the sentence fragment preceding the picture was roughly matched between the various experimental conditions (varying between 6.40 and 7.52 words on average across the eight different language × cognate status × sentence-constraint conditions). The sentences in the Dutch and English conditions were generally translations of one another, except when structural differences between Dutch and English led to awkward or incorrect translations, in which case for the other language condition a different sentence was constructed surrounding the critical target, but one that resembled the original sentence as closely as possible.

Table 3 shows some further examples of the sentences that were used. A complete list of the critical sentences can be obtained from the authors.

To check the constraint manipulation, a norming study was conducted. Thirty participants, all different from those participating in the experiment proper but drawn from the same population, were shown a booklet containing either the first part of all 100 Dutch sentences (17 participants) or the first part of all 100 English sentences (13 participants). The sentence parts in question contained all words preceding the target word (the elicited word) and ended with three full stops (e.g., "One usually eats soup with a . . . " or "She gave the guest a . . . "). The sentence parts relating to the different experimental conditions within each language condition (high and low-constraint, cognate and non-cognate) were presented in a randomized order, 10 on each page of the booklet, the 10 pages of a booklet being shuffled anew for every next participant. The participants were asked to fill in the missing word and a "predictability score" was calculated for each of the sentences. Two requirements for including a sentence in the actual experiment was that no other word than the intended target word received a high predictability score and that the scores for the missing words in the low-constraint sentences were all smaller than the score for the least predictable one of the missing words in the set of high-constraint sentences. Additional requirements were that within each language condition the cognate and non-cognate targets in the high-constraint condition were equally predictable and that the same held for the cognate and non-cognate targets in the low-constraint condition. A series of *t*-tests that was performed on the predictability scores confirmed that these criteria were met. Table 4 (on next page) shows the predictability scores for all experimental conditions.

In addition to the 100 critical sentences per language condition, 100 filler sentences were created for each language condition, two for each of the 50 filler pictures. The filler sentences had a structure similar to the critical

Table 4. *Mean predictability scores for the pictures' cognate and non-cognate names in both language conditions (critical stimuli only; standard deviations in parentheses).*

Condition	Cognates	Non-cognates
Dutch		
High-constraint	.91 (.13)	.90 (.17)
Low-constraint	.07 (.09)	.09 (.15)
English		
High-constraint	.76 (.30)	.84 (.17)
Low-constraint	.05 (.12)	.06 (.09)

sentences' structure, but the picture in a filler sentence did not always appear in the sentence's middle region but could also appear elsewhere in the sentence. The reason for this was to prevent the participants from anticipating the occurrence of the picture to be named while paying little attention to the remaining parts of the sentence. Spreading the presentation of the picture over different sentence regions would discourage such a strategy. When splitting up the filler sentences in three (about) equal parts, in 33, 36, and 31 of them the picture appeared in the sentence's first, second, and third part, respectively. This held for both the Dutch and the English filler sentences. As a further means to encourage the participants to pay attention to the sentence contexts, and to evaluate whether they indeed did do so, 20 yes/no comprehension questions regarding an immediately preceding filler sentence appeared on the screen at random positions across the experiment. The participants notified their answer to these questions by pushing a "yes" or "no" button.

Filler sentences were also used to allow the participants to recover from an error and to restore the normal processing routine following a question. A filler sentence was presented immediately after the participants had answered a question or made an error. Finally, filler sentences were used as warm-up trials (see Procedure below for details). As judged by the experimenters, with a few exceptions, the pictures in the filler sentences could not be predicted from prior context. Thus, only about 25% of the sentences presented to a participant were predictive of the embedded pictures.

Procedure

The participants were familiarized with the experimental materials in the same way as in Experiment 1. Next, the experiment proper started, in which a self-paced reading procedure was used. For each sentence, first a fixation point (a "+" sign) was presented in the center of the screen. When participants pressed the space bar, the fixation point was replaced by the first word of the sentence. When they

subsequently pressed the space bar again, the word on the screen was replaced by the next word of the sentence. This process continued until upon pressing the space bar the picture instead of a word appeared. The participant then named the picture in either Dutch or English, depending upon the language condition. The experimenter typed in a code indicating whether the response was correct or false or whether the voice key had malfunctioned. When an error was made, the experimenter provided the correct name to the participant. The typing of the code triggered the presentation of the next word of the sentence or, in case the picture concluded sentence presentation, the next trial. Between the presentations of two sentences an empty screen was presented for 1000 ms. Time duration between the onset of the picture and the moment the participant initiated the vocal response was registered as response time.

Like in Experiment 1, the actual experiment started with the presentation of 20 warm-up trials that allowed participants to get used to the task and the experimental procedure. Warm-up trials were randomly selected from the filler materials. Subsequently the critical and filler sentences, 200 in all, were presented in random order, using a different randomization for each participant. The stimuli were presented in eight groups with a short break after each group. Like in Experiment 1, an additional trial was presented after trials in which participants made an error or the voice key malfunctioned. Such additional trials used stimuli that were randomly selected from the filler materials. Data obtained for warm-up trials and additional trials were not recorded.

The comprehension questions consisted of yes/no questions and were presented in full in the center of the screen. Participants responded by typing the first letter of their response ("yes" or "no" in Condition English, "ja" or "nee" in Condition Dutch). When they gave the wrong answer, a blank screen was presented for 500 ms, after which the word "Wrong!" was presented for 1 s in the language of the condition. Then the next sentence was presented. Following the presentation of the last sentence, the exit questionnaire was administered. The whole experiment took about one hour to complete.

Results

Responses to the comprehension questions
The participants in the Dutch and English conditions gave correct answers to the comprehension questions about the sentences they had just read in 89.3% and 91.5% of the cases, respectively. We conclude that participants read the rebus sentences attentively and comprehensively.

Picture naming
Data treatment was identical to that in Experiment 1. The removal of incorrect responses, trials in which the voice

Table 5. *Participant mean reaction times (in milliseconds) per condition and error percentages (in parentheses) for Experiment 2.*

Condition	Cognates	Non-cognates	Cognate effect
Low-constraint			
Dutch	620 (2.4)	637 (2.1)	17°
English	758 (3.5)	802 (4.9)	44°;♦
Language effect	138	165	
High-constraint			
Dutch	475 (0.9)	452 (0.9)	–23
English	599 (3.3)	627 (4.2)	28•
Language effect	124	175	

°$p < .01$ in the subject analysis; •$p < .05$ in the subject analysis; ♦$p < .05$ in the item analysis

Note: High-constraint = the sentence context is predictive of the upcoming picture; low-constraint = the sentence context is not predictive of the upcoming picture; Cognate effect = Non-cognates RT – Cognates RT. Language effect = English RT – Dutch RT. The significance level for the cognate effects was calculated using one-sided *t*-tests.

key malfunctioned, extremes, and RTs that deviated more than 3 *SD*s from the participant cell means, accounted for 1.6%, 1.7%, 0.7%, and 1.5% of the data, respectively, in Condition Dutch, and for 4.0%, 1.3%, 0.5%, and 1.4% of the data, respectively, in Condition English. The remaining RTs were used in the calculation of the means for the cognate and non-cognate pictures separately in each sentence-constraint by language condition. These means are presented in Table 5, as are the error scores for the various conditions.

An analysis of variance (ANOVA) was performed on these means per participant per condition, with language (two levels: Dutch and English) as between-subjects variable and cognate status (two levels: cognates and non-cognates) and sentence constraint (two levels: high and low) as within-subjects variables. The corresponding item analysis was also performed, with language (two levels: Dutch and English) and cognate status (two levels: cognates and non-cognates) as between-items variables and sentence constraint (two levels: high and low) as within-items variable. Note that, just as in the analyses of Experiment 1, the power to detect a cognate effect is much larger in the subject analysis than in the item analysis.

The effect of language was significant, $F_1(1,44) = 38.62$, $MSE = 27,058$, $p < .001$, and $F_2(1,96) = 169.93$, $MSE = 6,667$, $p < .001$. Participants in Condition English responded more slowly ($M = 697$ ms) than those in Condition Dutch ($M = 546$ ms). The effect of sentence constraint was significant, $F_1(1,44) = 492.75$, $MSE = 2,565$, $p < .001$, and $F_2(1,96) = 474.23$, $MSE = 2,981$, $p < .001$. Participants responded faster in the high-constraint condition ($M = 538$ ms) than in the low-constraint condition ($M = 704$ ms). The effect of cognate status was

significant in the participant analysis, $F_1(1,44) = 8.50$, $MSE = 1,530$, $p = .006$, but not in the item analysis, $F_2(1,96) = 2.15$, $MSE = 6,667$, $p = .15$. Participants responded faster in the cognate condition ($M = 613$ ms) than in the non-cognate condition ($M = 630$ ms). The interaction between cognate status and sentence constraint was significant in the participant analysis, $F_1(1,44) = 7.75$, $MSE = 1,173$, $p = .008$ and almost significant in the item analysis, $F_2(1,96) = 3.75$, $MSE = 2,981$, $p = .056$. The combined cognate effect for the Dutch and English conditions was absent in the high-constraint condition ($M = 3$ ms) and present in the low-constraint condition ($M = 31$ ms). The interaction between cognate status and language was significant in the participant analysis, $F_1(1,44) = 11.71$, $MSE = 1,530$, $p = .001$, and showed a trend towards significance in the item analysis, $F_2(1,96) = 2.79$, $MSE = 6,667$, $p = .098$. The combined cognate effect for the high-constraint and low-constraint condition was absent in Dutch ($M = –3$ ms) and present in English ($M = 37$ ms). The interaction between sentence constraint and language was not significant, both Fs < 1. The triple interaction between cognate status, language and sentence constraint was not significant, both ps > .25.[7]

To evaluate the significance of the simple effects of cognate status, we calculated the significance of the cognate effect for each language and each constraint condition separately by using planned comparisons that consisted of one-sided paired (in the subject analysis) and unpaired (in the item analysis) *t*-tests. The results of these analyses are summarized in Table 5. In the participant analysis, the cognate effect was present in three of the four conditions, the exception being the high-constraint condition in Dutch. A post-hoc two-sided paired *t*-test showed that in this condition cognates were in fact responded to more slowly than non-cognates ($p = .016$), but we believe that this latter finding should not be attributed to our cognate manipulation.[8] In the item analysis, the cognate effect only surfaced in the

---

[7] The participants named each picture twice, once in the high-constraint condition and a second time in the low-constraint condition. However, task practice may reduce the cognate effect (see footnote 5 above). To see whether such reduction might have occurred in Experiment 2 we performed a further pair of analyses (by participants and by items) that, in addition to the variables language, cognate status, and sentence constraint, included the new variable presentation order (first vs. second). The main effect of presentation order was significant: Pictures presented for the first time were named more slowly (664 ms) than those presented for the second time (582 ms), $F_1(1,44) = 159.15$, $MSE = 3,899$, $p < .001$, and $F_2(1,96) = 209.40$, $MSE = 3,597$, $p < .001$. The interaction between presentation order and cognate status was, however, not significant, $F_1(1,44) = 1.52$, $MSE = 2,550$, $p = .22$, and $F_2 < 1$, indicating that task practice had not diminished the cognate effect.

[8] Given the overall very short response times in the Dutch high-constraint condition, just a couple of outlier items in this condition might have been responsible for this unexpected outcome. Inspection of the data showed that two cognate items in this condition produced

low-constraint English condition. Finally, we computed the size of the cognate effect in each constraint condition for each participant of both language conditions. Paired *t*-tests showed that the size of the cognate effects in the low-constraint and high-constraint conditions did not statistically differ when participants named the pictures in English ($p > .35$). In contrast, the size of the obtained cognate effects in the low-constraint and high-constraint conditions did differ when participants named the pictures in Dutch, $t(22) = 3.81$, $p = .001$.

To see whether speed–accuracy trade-offs occurred, the corresponding error analyses were also performed. They showed an effect of language, $F_1(1,44) = 13.00$, $MSE = 1.27$, $p = .001$, and $F_2(1,96) = 10.84$, $MSE = 1.40$, $p = .001$. More errors were made in Condition English ($M = 4.0\%$) than in Condition Dutch ($M = 1.6\%$). The effect of sentence constraint was significant in the participant analysis, $F_1(1,44) = 4.86$, $MSE = 0.49$, $p = .033$, but not in the item analysis, $F_2(1,96) = 2.06$, $MSE = 1.07$, $p = .15$. More errors were made in the low-constraint condition ($M = 3.2\%$) than in the high-constraint condition ($M = 2.3\%$). None of the other main effects and interactions was significant. We conclude that speed–accuracy trade-offs were absent in the data.

Discussion

We found an effect of cognate status when Dutch–English bilingual participants named pictures in a sentence context in L2 English: The participants named pictures that had Dutch–English cognate names faster than pictures with Dutch–English non-cognate names. Whether or not the lead-in sentence fragments were predictive of the upcoming pictures influenced the overall reaction time but did not influence the size of the cognate effect in Condition English: The cognate effect was equally large statistically in the high- and low-constraint conditions. In contrast, when the pictures appeared in an L1 Dutch context and were named in Dutch, the common cognate effect only occurred in the low-constraint condition, where the

---

exceptionally long mean RTs (of 597 ms and 574 ms), which were longer than ALL RTs for non-cognate items in this condition. Inspection of the corresponding high-constraint sentences showed that one sentence contained an ambiguous word (*roken*, which can denote both "to smoke" and "smelled" in Dutch), which may have led some of our participants down a garden path, delaying the response. The other sentence also contained an ambiguous word (*huilen*, which can denote both "to cry" and "to howl" in Dutch, "to cry" being the most frequent meaning whereas "to howl" was the intended meaning in the sentence), which may have activated the wrong target name (*baby* instead of *wolf*). When the RTs associated with these two deviant items were excluded from the analysis, the difference between the RTs for cognate and non-cognate pictures in the Dutch high-constraint condition (465 ms vs. 452 ms) was no longer significant ($p = .11$). We therefore treat the reversed cognate effect in the Dutch high-constraint condition that showed post-hoc significance as a fluke.

lead-in sentence fragments were not predictive of the upcoming pictures.

Experiments 1 and 2 combined

To evaluate the influence of language and sentence context on the cognate effect, part of the data of Experiments 1 and 2 were aggregated. Because in Experiment 1 the picture was presented in isolation and, thus, could not be predicted by the participants, we only used the data of the low-constraint (non-predictive) condition from Experiment 2, this condition being most comparable to Experiment 1. The ensuing data were analyzed with an ANOVA with cognate status (two levels: cognates and non-cognates) as a within-subjects variable and language (two levels: Dutch and English) and context (two levels: with and without) as between-subjects variables. The corresponding item analysis was also performed, with cognate status and language as between-items variables and context as within-items variable.[9]

The effect of cognate status was significant, $F_1(1,88) = 100.40$, $MSE = 1,752$, $p < .001$; $F_2(1,96) = 17.44$, $MSE = 10,176$, $p < .001$. Participants named pictures with cognate names ($M = 736$) faster than pictures with non-cognate names ($M = 798$). The effect of language was significant, $F_1(1,88) = 60.79$, $MSE = 21,737$, $p < .001$, and $F_2(1,96) = 130.28$, $MSE = 10,176$, $p < .001$. Participants were faster in Dutch ($M = 682$) than in English ($M = 852$). The effect of context was significant, $F_1(1,88) = 33.46$, $MSE = 21,737$, $p < .001$, and $F_2(1,96) = 467.84$, $MSE = 1,499$, $p < .001$. Participants were slower naming pictures in isolation ($M = 830$) than in a sentence context ($M = 704$). The interaction between language and context was not significant in the participant analysis ($F < 1$) but showed a trend towards significance in the item analysis ($p = .056$). The interactions between cognate status and context and between cognate status and language were significant, all $ps < .023$, but qualified by a triple interaction between cognate status, language, and context, $F_1(1,88) = 12.54$, $MSE = 1,752$, $p = .001$, and $F_2(1,96) = 11.40$, $MSE = 1,499$, $p = .001$. The triple interaction was further examined by computing the size of the cognate effect in both language conditions in each of the two experiments for each participant. An ANOVA was

---

9 Our cognate materials showed variation in cognate similarity. To evaluate the influence of this variable, we reran these analyses using only the cognate materials and replacing the cognate status variable with a cognate similarity variable. Cognate similarity (high or low) was determined for each picture name by dividing our cognate materials in two sets, one with the 12 items with the highest cognate similarity (as calculated using Van Orden's formulae, see Appendix A for details) and one with 12 remaining items (the item with a median cognate similarity score was discarded in these analyses). The results showed that pictures with more similar cognate names contributed more to the overall cognate effects than pictures with less similar ones.

performed on these effect sizes with condition (four levels: Dutch naming in isolation, English naming in isolation, Dutch naming in context, English naming in context) as a between-subjects variable. The effect of condition was significant, $F(3,88) = 23.77$, $MSE = 3,505$, $p < .001$. Dunnett T3 post-hoc analysis showed that the size of the cognate effect when pictures were named in isolation in English was larger (151 ms) than the sizes of the cognate effect in all other conditions (between 17 ms and 44 ms), all ps $< .001$. The effect sizes of the other conditions did not differ from each other, all ps $> .32$.

In conclusion, pictures were named faster in L1 Dutch than in L2 English; for each language, pictures were named faster when presented in a low-constraint context than when presented in isolation; pictures with cognate names were named faster than pictures with non-cognate names, and this cognate effect was largest when the pictures were presented in isolation and had to be named in L2 English.

General discussion

Word production in isolation

In Experiment 1 we examined whether the production of isolated words by bilinguals is language-selective or language-nonselective, and whether relative proficiency in the two languages might modulate any effect the non-target language may have on word production in the target language. The cognate status of the pictures' names was manipulated and the cognate effect – shorter response times for pictures with cognate names than for pictures with non-cognate names – was considered the critical marker of language-nonselective word production. This effect occurred both when the participants named the pictures in L1 Dutch and when they named them in L2 English, but it was considerably larger in L2 English. This latter finding indicates that the influence of the non-target language on word production in the target language is modulated by relative language proficiency. The results replicate similar findings in the literature (e.g., Christoffels et al., 2006; Costa et al., 2000; Hoshino & Kroll, 2007; Poarch & Van Hell, 2012; see the introduction for a discussion of other related studies).

Explaining the cognate effect
How can the cognate effects obtained with word production in isolation and their modulation by relative language proficiency be explained? If it is assumed that activation processes during word production are language-nonselective all the way down to sublexical phonology, the explanation of the cognate effect is rather straightforward. A common account, first suggested by Costa et al. (2000), was already introduced and illustrated before (see the introduction, Figure 1; see also Dell, Schwartz, Martin,

Saffran & Gagnon, 1997). It assumes that the processing of a picture involves the activation of first, a set of semantic nodes, and second, the corresponding lexical nodes in BOTH languages. All activated lexical nodes propagate activation forward to the sublexical layer. As a result, the sublexical phonological nodes that a cognate shares between languages (e.g., English *mouse* and Dutch *muis*; Figure 1, left) receive more activation than the non-shared sublexical phonological nodes of a non-cognate (e.g., English *axe* and Dutch *bijl*; Figure 1, right). In addition, due to feedback in the model from the phonological layer back to the lexical layer, also the lexical nodes of cognates receive more activation than the lexical nodes of non-cognates, activation which is then fed forward to the corresponding sublexical nodes again. Consequently, naming is relatively fast for pictures with cognate names.

If, however, activation during picture naming were language-selective, pictures with cognate names and those with non-cognate names should be named equally rapidly because the sublexical phonological nodes of both cognates and non-cognates would receive activation from one source only. The very occurrence of the cognate effect thus suggests that bilingual word production in isolation is language-nonselective.

The present finding that the cognate effect is so much larger when pictures are named in the weaker language L2 than when named in the stronger language L1 can be explained by assuming that for our participants the connections between the semantic and lexical nodes and between the lexical and sublexical nodes were stronger for L1 than for L2 (e.g., Costa et al., 2000; Kroll & Stewart, 1994; cf., Jescheniak & Levelt, 1994; weaker connections are represented by dotted lines in Figure 1). Because stronger links transmit more activation than weaker links, the retrieval of a cognate's name in L2 benefits more from activation in L1 than vice versa. The likely cause for such a between-language difference in the strength of these connections is a difference in usage frequency between the two languages: The more often a particular memory connection is exploited, the stronger it gets.

Word production in sentence context

In Experiment 2 we examined whether sentence context modulates the cognate effect and whether the degree to which prior sentence context constrains the set of possible picture names might modulate the cognate effect. Like in Experiment 1, the cognate effect was relatively large when the pictures were named in L2 English. Additional findings were that three of the four sentence-constraint by language conditions showed a significant cognate effect, the exception being the high-constraint/Dutch condition. Furthermore, an overall analysis of the cognate effects

as observed in the English and Dutch conditions of Experiment 1 (with pictures presented in isolation) and those observed in the English and Dutch low-constraint condition of Experiment 2 showed that the effect was exceptionally large when pictures were named in English and, at the same time, in isolation. In the remaining three conditions the effect was statistically equally large. Finally, sentence context in general accelerated the responses considerably and this accelerative effect of sentence context was exceptionally large in the high-constraint sentence condition.

These findings extend the results from earlier studies on bilingual visual word recognition in context to bilingual word production in context. These earlier studies had shown that the cognate effect in word recognition survives in a low-constraint context, both when the task had to be performed in L2 (Duyck et al., 2007; Schwartz & Kroll, 2006; Van Hell & De Groot, 2008) and when task performance was in L1 (Titone et al., 2011; Van Assche et al., 2009), but that it does not survive in a high-constraint context (Schwartz & Kroll, 2006; Van Hell & De Groot, 2008) or only very early on during L2 word processing (Libben & Titone, 2009, but see Van Assche et al., 2011). The present study indicates that the cognate effect in word production is also affected by sentence constraint because the usual cognate effect (shorter RTs for cognates) disappeared in the high-constraint L1 Dutch condition. Generally then, both the recognition studies and the present production study show that degree of sentence constraint affects the non-target language's influence on processing.

Finally, the present results agree with the word-translation study of Van Hell and De Groot (2008) that, as mentioned in the introduction, is most similar to the present study. The translation task in Van Hell and De Groot's study also showed an effect of both context and sentence constraint. An apparent difference between the results of the Van Hell and De Groot study and the present one is that the cognate effects appear to be larger in the former study. A plausible reason for that difference (which, due to design differences between the two studies, cannot be tested statistically) is that in word translation, in addition to facilitated name retrieval for cognates due to shared phonological sublexical OUTPUT nodes, two further sources for the cognate effect exist, both related to INPUT processing prior to conceptual activation (see De Groot, 2011, pp. 261–265). In addition, it might be argued that cognates presented for translation activate the sublexical phonological nodes of their translation equivalents directly, by means of grapheme-to-phoneme connections. This latter explanation is in accordance with the standard explanation of the effects of phonological and/or orthographic overlap between a distractor and a picture's name in picture-naming studies that employ the picture–word task (see e.g., Starreveld, 2000).

Explaining influences of sentence context on the cognate effect

How can the effects of context (with or without), sentence constraint, and language on the cognate effect be explained? We suggest an interpretation in terms of two causes of differential activation in the memory structures illustrated in Figure 1 during word production in isolation and word production in context. First, a sentence context boosts the activation of all elements in the target language system as compared with the in-isolation condition. This boost may stem from the appropriate "language cue" in the conceptual message (La Heij, 2005; see also Poulisse & Bongaerts, 1994) being activated more when pictures are presented in context than when they are presented in isolation. The heightened activation of the language cue then causes the target language system as a whole to be relatively highly activated. Second, prior sentence context pre-activates a part, or all, of the upcoming picture's semantic representation. In a high-constraint sentence context the amount of pre-activation might even be comparable to the activation caused by the presentation of the picture itself.

The first process (additional activation of the complete target-language system) generally leads to a reduction of the cognate effect. It also explains that the overall response times are shorter for picture naming in context than for picture naming in isolation. The second process (the pre-activation of the upcoming picture's semantic representation) reduces the cognate effect only under specific circumstances, namely, when the pre-activation is so strong that the picture's name can be predicted in advance of the presentation of the picture. Under these circumstances, a cognate effect might in fact even fail to surface altogether in the response times, as we explain below. This second process also accounts for the fact that the overall response times are shorter in the high-constraint context condition than in the low-constraint context condition. Importantly, the reduction of the cognate effect due to these two processes occurs while the bilingual word-production system still operates in a language-nonselective manner.

To clarify these statements, the notion of iterations as assumed in neural network models of word production (Dell, 1986; Levelt, Roelofs & Meyer, 1999; Starreveld & La Heij, 1996), and how they relate to response time, is relevant. For present purposes we assume that the production of a name depends on the moment the activation level in the representation of the picture's name reaches a critical threshold level. We call this the moment of selection. Selection is performed after a number of processing steps (called "iterations") during each of which a certain amount of activation is fed through the network. If selecting a non-cognate name takes x iterations and selecting a cognate name takes y iterations, the model would generate a cognate effect of $x - y$ iterations (each

iteration representing z ms; for example, $z = 25$ in the model of Levelt et al. (1999), and $z = 10$ in the model of Starreveld & La Heij (1996)).

Importantly, such neural networks possess the feature that increasing the amount of activation in the network results in a decrease of the number of iterations required to select a target node. For models that use linear activation functions (e.g., Dell, 1986; Levelt et al., 1999), the amount of activation is inversely proportional to selection time. For instance, when the activation that reaches a node is doubled, the model would need half of the original number of iterations to select that node. In our example, doubling the amount of activation that is transmitted per iteration has the effect that $\frac{1}{2}x$ and $\frac{1}{2}y$ iterations are required to select a non-cognate name and a cognate name, respectively; the cognate effect would thus be reduced from $x - y$ to $\frac{1}{2}(x - y)$. Therefore, if sentence context indeed has the effect of increasing the overall activation in the target language system, smaller cognate effects should result when word production is performed in sentence context than when performed in isolation. In addition, the fact that fewer iterations are required to reach the critical threshold of activation when pictures are named in a sentence context than when they are named in isolation causes the responses to be faster in the former case. As we have seen, both these effects indeed occurred.[10]

To account for the effects of sentence constraint (high vs. low), we may turn to the view that the processing of sentence context leads to the pre-activation of semantic features that are part of the semantic representation of the upcoming target. This view emerged from a number of studies on visual word recognition, which employed various paradigms including RT measurements, eye movement registration, and ERP measurements (e.g., Altarriba, Kroll, Sholl & Rayner, 1996; Dahan & Tanenhaus, 2004; Otten, Nieuwland & Van Berkum, 2007; Schwanenflugel & LaCount, 1988; Schwanenflugel & Shoben, 1985; Van Berkum, Brown, Zwitserlood, Kooijman & Hagoort, 2005). Schwanenflugel and Shoben (1985) and Schwanenflugel and LaCount (1988) distinguish between two types of semantic features activated by sentence context, specific and more general

ones. The authors hypothesized that low-constraint sentence contexts lead to the activation of fewer and more general features than high-constraint sentence context. For example, a low-constraint sentence context like "The lady was a competent ..." might only activate the general features [skilled activity] and [something humans can be], whereas a high-constraint context fragment like "John kept his gym clothes in a ..." might, in addition to some general features, also activate the specific semantic features [small], [rectangular], [associated with gyms], [holds clothes], and ["shutable"], which together form the distributed semantic representation of a locker. (The example is from Schwanenflugel & Shoben, 1985.)

On this account, a plausible reason why the cognate effect might disappear in a high-constraint sentence context is that such contexts lead to the full activation of semantic nodes that are part of the upcoming picture's semantic representation. In cases where all of the semantic nodes representing a picture's meaning are fully and rapidly activated this way, the exact name of the picture might be available prior to the picture's occurrence (i.e., the name can be predicted). Since this process of pre-activation applies to pictures with cognate and non-cognate names alike, a null-effect of cognate status would result. Phrased in terms of the neural-network account presented above, both for cognate names and non-cognate names a sufficiently large number of iterations may have ensued to reach the activation level required for selection even before the target picture was actually presented. This situation may have arisen in the Dutch high-constraint condition. The process of full pre-activation also accounts for the extremely fast responses that we observed in this condition.

In conclusion, it appears that language users can and do predict upcoming words from semantically highly-constraining contexts. The results of the present study suggest that this prediction process is also operative with subsequent word production instead of word recognition. But why did a cognate effect still turn up in the English high-constraint condition even though both cognate names and non-cognate names were pre-activated by the sentence context? The reason may lie in our participants' relatively low proficiency in L2 English (as compared with L1 Dutch) and the consequent slower activation of the relevant semantic features and, next, of the corresponding lexical and sublexical nodes. As a result, by the time the picture is presented the prediction of the target name may not be fully enabled yet, and a cognate effect shows up. Alternatively, it might have been the case that in a proportion of the trials the picture names could be predicted whereas in the remainder of the trials such prediction was not fully enabled yet. The latter situation will more often hold for non-cognate names than for cognate names (the former requiring more iterations to reach their threshold level of activation) and a reduced cognate effect shows up.

---

[10] An alternative account of the effect of sentence context might be developed based on the idea that, given dual language activation, the selection of the picture's name involves the inhibition of potential competitors belonging to the non-target language, especially the name of the picture in the non-target language (e.g., Green, 1998). In such an account, sentence context might increase the degree of inhibition of the non-target language. As a result, the lexical node of a cognate in the non-target language is activated less and thus sends on less activation to the sub-lexical nodes that the cognate shares between languages. Consequently, the cognate effect (the difference in the number of iterations to reach the critical threshold level of activation for naming cognates and non-cognates) will be smaller when pictures are named in a sentence context than when they are named in isolation.

Finally, in the low-constraint condition full prediction plausibly never occurred (neither in Condition Dutch nor in Condition English; neither for cognates nor for non-cognates) for reasons that the sentence context will pre-activate only a few, rather general, semantic features. What this type of pre-activation will do though is give the naming of both cognate names and non-cognate names a head start as compared with naming the pictures in isolation. However, as explained above, the main cause for faster selection in this condition is the general increase in activation of the complete target language system induced by the sentence context.

Summary and conclusion

To summarize, we hypothesize that the cognate effects and response rates obtained in the various conditions of this study can be explained in terms of a combination of the following processes and characteristics of the memory structures involved in bilingual word production: (i) Facilitated retrieval of cognate names due to extra activation in the relevant sublexical phonological output nodes which they receive from the corresponding non-target language's lexical nodes; (ii) Stronger connections between the semantic, lexical, and sublexical nodes in dominant L1 than in weaker L2; (iii) A general increase of activation in the target language system caused by sentence context; and (iv) The pre-activation of a part, or all, of the upcoming picture's semantic representation caused by sentence context. Importantly, we related these processes to implemented neural network models of monolingual language production.

The overarching goal of this study was to fill in a gap in the current literature concerning the question whether language processing in bilinguals is language-nonselective by examining cognate effects during word production in sentence context. The general pattern of results indicates that the bilingual language-processing system operates in a fundamentally language-nonselective way.

Appendix A. Van Orden's (1987) formulae

Van Orden's formula for graphemic similarity (GS) is: GS = 10([[50F + 30V + 10C)/A] + 5T + 27B +18E). F = number of pairs of adjacent letters in the same order shared by word pairs; V = number of pairs of adjacent letters shared by word pairs but in reverse order; C = number of single letters shared by word pairs; A = average number of letters in the two words; T = ratio of number of letters in the shorter word to the number of letters in the longer word; B = 1 if the two words have the same word-initial letter; if not, then B = 0. Finally, E = 1 if the two words have the same word-final letter; if not, then E = 0. Orthographic similarity (OS) between word X and Y is OS$_{XY =}$ GS$_{XY}$/GS$_{YY.}$ See Van Orden (1987) for further details and examples.

Appendix B. English and Dutch Names of the Critical Pictures in Experiments 1 and 2

Phonetic transcriptions are given in parentheses.

Cognates

Anchor–anker (æŋkər–ɑŋkər); apple–appel (æpəl–ɑpəl); baby–baby (beɪbi–bebi); beard–baard (bɪərd–bart); boat–boot (boʊt–bot); book–boek (bʊk–buk); bread–brood (brɛd–brot); clock–klok (klɒk–klɔk); cross–kruis (krɒs–krœys); door–deur (dɔr–dʌr); flag–vlag (flæg–vlɑx); foot–voet (fʊt–vut); hand–hand (hænd–hɑnt); heart–hart (hɑrt–hɑrt); house–huis (haʊs–hœys); lamp–lamp (læmp–lɑmp); mouse–muis (maʊs– mœys); pear–peer (pɛər–pɪr); pipe–pijp (paɪp–pɛip); ring–ring (rɪŋ–rɪŋ); tank–tank (tæŋk–tɛŋk); tent–tent (tɛnt–tɛnt); train–trein (treɪn–trɛin); violin–viool (vaɪəlɪn–vi'ɔl); wolf–wolf (wʊlf–ʋɔlf).

Non-cognates

Ax–bijl (æks–bɛil); basket–mand (bæskɪt–mɑnt); bicycle–fiets (baɪsɪkəl–fits); bottle–fles (bɒtl–flɛs); box–doos (bɒks–dos); cheese–kaas (tʃiz–kas); church–kerk (tʃɜrtʃ–kɛrk); cloud–wolk (klaʊd–ʋɔlk); couch–bank (kaʊtʃ–bɑnk); deer–hert (dɪər–hɛrt); dog–hond (dɒg–hɔnt); fly–vlieg (flaɪ–vlix); glasses–bril (glæsəs–brɪl); knife–mes (naɪf–mɛs); letter–brief (lɛtər–brif); mountain–berg (maʊntn–bɛrx); shark–haai (ʃɑrk–haj); spoon–lepel (spun–lepəl); stairs–trap (stɛərz–trɑp); tear–traan (tɪər–tran); toe–teen (toʊ–ten); tree–boom (tri–bom); witch–heks (wɪtʃ–hɛks); woman–vrouw (wʊmən–vrʌu); zipper–rits (zɪpər–rɪts).

References

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers, 36,* 140–155.

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. *Memory & Cognition, 24,* 477–492.

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

Blumenfeld, H. K., & Marian, V. (2007). Constraints on parallel activation in bilingual spoken language processing: Examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes, 22,* 633–660.

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23,* 45–93.

Brown, A. S. (1979). Priming effects in semantic memory retrieval processes. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 65–77.

Chambers, C. G., & Cooke, H. (2009). Lexical competition during second-language listening: Sentence context, but

not proficiency, constrains interference from the native language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 1029–1040.

Christoffels, I. K., De Groot, A. M. B., & Kroll, J. F. (2006). Memory and language skills in simultaneous interpreters: The role of expertise and language proficiency. *Journal of Memory and Language, 54,* 324–345.

Colomé, A. (2001). Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language, 45,* 721–736.

Colomé, A., & Miozzo, M. (2010). Which words are activated during bilingual word production? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 96–109.

Costa, A., Caramazza, A., & Sebastián-Gallés, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1283–1296.

Costa, A., Colomé, A., Gómez, O., & Sebastián-Gallés, N. (2003). Another look at cross-language competition in bilingual speech production: Lexical and phonological factors. *Bilingualism: Language and Cognition, 6,* 167–179.

Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language, 41,* 365–397.

Costa, A., & Santesteban, M. (2004). Bilingual word perception and production: Two sides of the same coin? *Trends in Cognitive Science, 8,* 253.

Dahan, D., & Tanenhaus, M. K. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 498–513.

De Groot, A. M. B. (2011). *Language and cognition in bilinguals and multilinguals: An introduction*. New York & Hove: Psychology Press.

De Groot, A. M. B., Borgwaldt, S., Bos, M., & Van den Eijnden, E. (2002). Lexical decision and word naming in bilinguals: Language effects and task effects. *Journal of Memory and Language, 47,* 91–124.

De Groot, A. M. B., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory. *The Quarterly Journal of Experimental Psychology, 53A,* 397–428.

De Groot, A. M. B., & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30,* 90–123.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review, 93,* 283–321.

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review, 104,* 801–838.

Dijkstra, T., De Bruijn, E., Schriefers, H., & Ten Brinke, S. (2000). More on interlingual homograph recognition: Language intermixing versus explicitness of instruction. *Bilingualism: Language and Cognition, 3,* 69–78.

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

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,* 175–197.

Dijkstra, T., Van Jaarsveld, H., & Ten Brinke, S. (1998). Interlingual homograph recognition: Effects of task demands and language intermixing. *Bilingualism: Language and Cognition, 1,* 51–66.

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. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 663–679.

Elston-Güttler, K. E., Gunter, T. C., & Kotz, S. A. (2005). Zooming into L2: Global language context and adjustment affect processing of interlingual homographs in sentences. *Cognitive Brain Research, 25,* 57–70.

Glaser, W. R., & Düngelhoff, F.-J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 640–654.

Gollan, T. H., Forster, K. I., & Frost, R. (1997). Translation priming with different scripts: Masked priming with cognates and noncognates in Hebrew–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1122–1139.

Grainger, J., & Dijkstra, T. (1992). On the representation and use of language information in bilinguals. In R. J. Harris (ed.), *Cognitive processing in bilinguals*, pp. 207–220. Amsterdam: Elsevier.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition, 1,* 67–81.

Grosjean, F. (2001). The bilingual's language modes. In J. Nicol (ed.), *One mind, two languages: Bilingual language processing*, pp. 1–22. Oxford: Blackwell.

Hermans, D., Bongaerts, T., De Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition, 1,* 213–229.

Hermans, D., Ormel, E., Van Besselaar, R., & Van Hell, J. [G.] (2011). Lexical activation in bilinguals' speech production is dynamic: How language ambiguous words can affect cross-language activation. *Language and Cognitive Processes, 26,* 1687–1709.

Hoshino, N., & Kroll, J. F. (2007). Cognate effects in picture naming: Does cross-language activation survive a change of script? *Cognition, 106,* 501–511.

Jared, D., & Kroll, J. F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language, 44,* 2–31.

Jared, D., & Szucs, C. (2002). Phonological activation in bilinguals: Evidence from interlingual homograph naming. *Bilingualism: Language and Cognition, 5,* 225–239.

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20,* 824–843.

Ju, M., & Luce, P. A. (2004). Falling on sensitive ears: Constraints on bilingual lexical activation. *Psychological Science, 15,* 314–318.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33,* 149–174.

La Heij, W. (2005). Selection processes in monolingual and bilingual access. In. J. F. Kroll & A. M. B. de Groot (eds.), *Handbook of bilingualism: Psycholinguistic approaches*, pp. 289–307. New York: Oxford University Press.

La Heij, W., De Bruyn, E., Elens, E., Hartsuiker, R., Helaha, D., & Van Schelven, L. (1990). Orthographic facilitation and categorical interference in a word-translation variant of the Stroop task. *Canadian Journal of Psychology, 44,* 76–83.

La Heij, W., Starreveld, P. A., & Steehouwer, L. C. (1993). Semantic interference and orthographic facilitation in definition naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 352–368.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–75.

Libben, M. R., & Titone, D. A. (2009). Bilingual lexical access in context: Evidence from eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 381–390.

Marian, V., & Spivey, M. (2003a). Bilingual and monolingual processing of competing lexical items. *Applied Psycholinguistics, 24,* 173–193.

Marian, V., & Spivey, M. (2003b). Competing activation in bilingual language processing: Within- and between-language competition. *Bilingualism: Language and Cognition, 6,* 97–115.

Miller, N. A., & Kroll, J. F. (2002). Stroop effects in bilingual translation. *Memory & Cognition, 30,* 614–628.

Otten, M., Nieuwland, M. S., & Van Berkum, J. J. A. (2007). Great expectations: Specific lexical anticipation influences processing of spoken language. *BMC Neuroscience, 8,* 89.

Poarch, G. J., & Van Hell, J. G. (2012). Cross-language activation in children's speech production: Evidence from second language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology, 111,* 419–438.

Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E., & Sherman, J. (1986). Pictures in sentences: Understanding without words. *Journal of Experimental Psychology: General, 115,* 281–294.

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics, 15,* 36–57.

Rodriguez-Fornells, A., Van der Lugt, A., Rotte, M., Britti, B., Heinze, H.-J., & Münte, T. F. (2005). Second language interferes with word production in fluent bilinguals: Brain potential and functional imaging evidence. *Journal of Cognitive Neuroscience, 17,* 422–433.

Sánchez-Casas, R. M., Davis, C. M., & García-Albea, J. E. (1992). Bilingual lexical processing: Exploring the cognate/non-cognate distinction. *European Journal of Cognitive Psychology, 4,* 293–310.

Schwanenflugel, P. J., & LaCount, K. L. (1988). Semantic relatedness and the scope of facilitation for upcoming words in sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14,* 344–354.

Schwanenflugel, P. J., & Shoben, E. J. (1985). The influence of sentence constraint on the scope of facilitation for upcoming words. *Journal of Memory and Language, 24,* 232–252.

Schwartz, A. I., & Arêas Da Luz Fontes, A. B. (2008). Cross-language mediated priming: Effects of context and lexical relationship. *Bilingualism: Language and Cognition, 11,* 95–110.

Schwartz, A. I., & Kroll, J. F. (2006). Bilingual lexical activation in sentence context. *Journal of Memory and Language, 55,* 197–212.

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

Spivey, M. J., & Marian, V. (1999). Cross talk between native and second languages: Partial activation of an irrelevant lexicon. *Psychological Science, 10,* 281–284.

Starreveld, P. A. (2000). On the interpretation of onsets of auditory context effects in word production. *Journal of Memory and Language, 42,* 497–525.

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 686–698.

Starreveld, P. A., & La Heij, W. (1996). Time course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 896–918.

Székely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., & Bates, E. (2003). Timed picture naming: Extended norms and validation against previous studies. *Behavior Research Methods, Instruments, & Computers, 35,* 621–633.

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. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 1412–1431.

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., & Hartsuiker, R. J. (2011). The influence of semantic constraints on bilingual word recognition during sentence reading. *Journal of Memory and Language, 64,* 88–107.

Van Assche, E., Duyck, W., Hartsuiker, R. J., & Diependaele, K. (2009). Does bilingualism change native-language reading? *Psychological Science, 20,* 923–927.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 443–467.

Van Hell, J. G., & De Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness

and cognate status in word association. *Bilingualism: Language and Cognition, 1,* 193–211.

Van Hell, J. G., & De Groot, A. M. B. (2008). Sentence context modulates visual word recognition and translation in bilinguals. *Acta Psychologica, 128,* 431–451.

Van Hell, J. G., & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review, 9,* 780–789.

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

Van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory & Cognition, 15,* 181–198.

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

Weber, A., & Cutler, A. (2004). Lexical competition in non-native spoken-word recognition. *Journal of Memory and Language, 50,* 1–25.