Visual cues for language selection in bilinguals

Robert J. Hartsuiker

Ghent University

Book chapter for Falk Huettig (Ed.). *Attention and Vision in Language Processing*. Springer.

Address for correspondence:

R. Hartsuiker

Department of Experimental Psychology

Ghent University

Henri Dunantlaan 2

B-9000 Ghent Belgium

Phone: +32 (0)9 2646 436

email: Robert.Hartsuiker@ugent.be

Abstract

Bilinguals need to select the right language for the particular context they are in, but how do they do this? One possibility is that they exploit visual cues from the context such as people's faces, so that recognition of the face increases the availability of the language associated with that face. This chapter first examines the degree to which bilinguals activate multiple languages and how this is constrained by *linguistic* cues and then discusses three new lines of research that investigate *visual* language cueing. Specifically, these new research lines suggest effects on language processing of (a) the language associated with familiar people's faces (e.g., German for Angela Merkel); (b) language that people associate with the race of unknown faces (e.g., Chinese for an Asian face, English for a white face); (c) language associated with cultural icons (e.g., Russian for a picture of the Kremlin).

Introduction

As an exchange student in the US, I once shared an apartment with a German-English bilingual; we consistently spoke English with each other. One day, however (in a state of considerable excitement about something) my roommate seemed to have forgotten which language we normally used, and produced a rather lengthy speech in German (until he was finally stopped by the look of surprise on my face). Such occurrences of language misselection seem to be rare however: we normally know exactly what language to speak at the very moment we see a familiar person. The question addressed in this chapter is whether bilingual speakers use information from the visual environment to help select the right language to speak or comprehend in. To set the stage, I first briefly review studies that have demonstrated parallel language activation in bilinguals and studies that have considered whether such parallel language activation can be modulated by linguistic cues.

1. Bilingual language processing is language non-selective.

Being able to restrict language processing to only the target language would seem to be very useful, given that there is a considerable amount of evidence that in lexical access, words from both a bilingual's languages become active (for recent reviews see Kroll & Bialystok, 2013). In the domain of *visual word recognition*, for instance, many studies have shown differences in the processing of words that are identical or similar in a bilingual's two languages (e.g., cognates such as Dutch and English *ring*, and interlingual homographs such as *list*, meaning trick in Dutch), as compared to words without any overlap in form. For instance, in a seminal study Dijkstra, Timmermans, and Schriefers (2000)

showed that Dutch-English bilinguals were much slower to respond that a string of letters was a Dutch word when it was an interlingual homograph between Dutch and English (as compared to a Dutch control word). What is more, interlingual homographs like *tree* that are high-frequent in English but low-frequent in Dutch were often not even accepted as Dutch words. Similar findings were obtained when the task was run in the subjects' second language, English. Similarly, studies with cognates, words that overlap both in form and meaning, have shown a cognate advantage in visual word recognition over control words. This cognate advantage does not only occur in subjecs' second language, but also in their first language (Van Hell & Dijkstra, 2002). This advantage is even larger when multilinguals have a third language for which the word is also a cognate (e.g., *echo* is a word in Dutch, English, and German, Lemhöfer, Dijkstra & Michel, 2004).

One might argue that such evidence is restricted to "special words" that belong to more than one language. But is it also the case that upon reading a language-unique word, bilinguals automatically activate the translation of that word in another language? Thierry and Wu (2007) recently provided evidence for the latter view. They asked monolingual speakers of English, monolingual speakers of Chinese, and Chinese-English bilinguals to judge the semantic relatedness of English word pairs such as *wife – husband* or *train - ham* (the Chinese monolinguals saw the Chinese translations of these pairs). These pairs were constructed so that in half of the experimental trials there was a shared character in the Chinese translations of the word pairs. For instance, the Chinese translation of *train* has the same initial character as that of *ham*. While

participants conducted this task, EEG was measured. As is to be expected, all groups showed a different brain potential as a function of whether the words were semantically related or not (i.e., an N400). But importantly, both the Chinese monolinguals (who saw the Chinese characters with or without orthographic overlap in between pair members) and the bilinguals (who only saw the words in English), showed differential brain potentials as a function of orthographic overlap (also in the N400 window, as well as in an earlier window) in Chinese (the English monolinguals did not show this component). These findings strongly suggest that upon reading a word in English, these bilinguals automatically activate the Chinese translation of that word up to the level of the orthographic code.

Evidence for language non-selective lexical access has also been found in other domains of language processing. In auditory word recognition, Spivey and Marian (1999) found that Russian-English bilinguals, immersed in an English-speaking environment, were sensitive to form similarity between a Russian word and the English name of an object. This was shown in a visual world eye-tracking experiment, in which subjects' eye movements to a visual scene where monitored as they listened to speech. Upon hearing a Russian instruction to move a stamp (*marku*), subjects were more likely to fixate a marker (related in phonological form) than an unrelated control object. Thus, even though the task was exclusively conducted in the subjects' native language (Russian), form-related words in their second language English became active, and did so to such an extent that it influenced looking behavior in a visual display.

Similarly, in spoken word production evidence for language non-selectivity has also been obtained. For instance, Colomè (2001) asked subjects to engage in a so-called phoneme monitoring task. On each trial, they were assigned a target phoneme (e.g., /m/) and saw a picture (e.g., of a table), and the task was to determine whether the Catalan name of the picture (*taula*) contained the target phoneme. The Spanish name (*mesa)* was an irrelevant dimension for the task, but nevertheless Catalan-Spanish bilinguals were slower to make a no-decision when the phoneme occurred in the Spanish name of the object (e.g., /m/ for *mesa*) compared to a phoneme that occurred in neither the Catalan nor Spanish name. This suggests that when producing the phonological code in Catalan, bilinguals also activate the phonological code in the non-target language, Spanish.

Summarizing, these and many other findings constitute clear evidence that bilinguals activate words from both of their languages during language processing and that this is the case irrespective of target language. One consequence is that bilinguals have more words to choose from, which may lead to a processing disadvantage. Indeed, a picture naming study (Ivanova & Costa, 2008) showed that monolingual speakers of Spanish were somewhat faster to name pictures of common objects than were Spanish-Catalan bilinguals, even though the bilinguals were using their first and dominant language. It might therefore be beneficial for bilinguals to restrict lexical access to the target language. This leads to the question of whether bilinguals can exploit cues, inherent in the linguistic signal or in the context in which that signal occurs, to "zoom into" (as Elston-Güttler, Gunter, & Kotz, 2005, called it) the right language.

One might argue that in Ivanova and Costa's study such cues were weak (as the stimuli to be named were not associated with a particular language and each word was produced as an isolated response, unaffected by any sentence context). We will now turn to the question of whether stronger language cues do effect the extent of language non-selectivity.

2. Do linguistic cues allow bilinguals to zoom into the right language?

The studies reviewed in the previous section have typically studied lexical access for single words that were read, or produced, without the larger context of a sentence or a discourse. One exception is the visual world study of Spivey and Marian, but in that experiment, the same sentence frame "now pick up the <object>" was repeated over and over again, rendering this different from more naturalistic sentence processing. It is very much conceivable that in the latter case, the language of the sentence provides an important cue about the language of each upcoming word, so that words from the other language no longer need to be considered. Consistent with this, De Bruijn, Dijkstra, Chwilla, and Schriefers (2001) gave example (1), which – to a Dutch-English bilingual – appears to be a perfectly good sentence in Dutch (although one with a rather unusual content). What is not obvious right away is that in fact every word of the sentence is an interlingual homograph with English.

(1) Door spot leek die brave, dove arts rover met pet

(because of mockery, that good, deaf doctor resembled robber with hat)

So does a sentence context help to rule out the irrelevant language? Duyck, Van Assche, Drieghe, and Hartsuiker (2007) argued that if it does, effects of cross-linguistic overlap (i.e., cognate status) should disappear once cognates are embedded in a sentence context (also see Schwartz & Kroll, 2006; Van Hell & De Groot, 2008). Duyck et al. first presented their subjects, Dutch-English bilinguals, with English-Dutch cognates and English control words in an English lexical decision experiment. They replicated the cognate effect (shorter reaction times for cognates than matched control words). Next, they embedded the cognates and control words in a sentence context; the last word of each sentence was always the cognate/control word, and participants made a lexical decision on the sentence-final word. Importantly, the cognate effect survived this manipulation. Interestingly, the cognate effect was stronger for a subset of the items that were completely identical in form between English and Dutch (*ring* – *ring*) than items that were similar, but not identical in form (*ship* – *schip*). A final experiment presented versions of the sentences (but now with critical items in an earlier sentence position, to avoid any effects of sentence wrap-up) and monitored subject's eye-movements. Both on relatively early measures (first fixation duration) and late measures (go-past time), there was an effect of cognate status, but this was the case only for the subset of identical cognates. Thus, in this experiment sentence context turned out to be insufficient to "turn off" the other language, but whether there is language non-selectivity under these conditions does seem to be modulated by degree of orthographic overlap.

Importantly, the Duyck et al. (2007) study was conducted in L2 English, and the subjects were clearly dominant in their L1 Dutch. One might argue that whereas one can never turn off a dominant L1 while processing L2, it should be possible to render L1 language non-selective. Van Assche, Duyck, Hartsuiker, and Drieghe (2009) therefore conducted an L1 version of the Duyck et al. study. They first established a cognate advantage with words presented in isolation (replicating Van Hell & Dijkstra, 2002). Importantly, in a sentence reading eye tracking experiment, there was also cognate facilitation, in the sense that the more similar a Dutch target word (e.g., *oven*) was to its English translation equivalent in orthographic and phonological form, the shorter various eye-movement measures were. A follow-up experiment replicated this finding with a further set of stimuli. Thus, Van Assche et al. concluded that learning a second language has a profound influence on how one reads text in the first language: if a bilingual reads her local newspaper in her native language, she does so differently from a monolingual.

The conclusion that language non-selectivity survives contextual cues is supported by studies in the domain of bilingual auditory word recognition. A language-ambiguous written word (e.g., *ring*) is completely identical in the two languages because the letters are identical in English and Dutch. In contrast, spoken words consist of a much richer signal that includes subphonemic (and perhaps prosodic) language information, which a listener might, in principle, exploit to restrict lexical access to the target language. For instance, Dutch and English use different allophones to realize the phone /r/; Spanish and English notoriously differ in the boundary between voiced and unvoiced consonants.

Lagrou, Hartsuiker, and Duyck (2011) therefore reasoned that if listeners can exploit such cues, and can zoom into the correct language, there should be no difference in the time to recognize an interlingual homophone like /beI/ (Bay in English – Bij [bee] in Dutch) as compared to a monolingual control word. In contrast, lexical decision times to such interlingual homophones were longer than to control words, and this was true both in tasks conducted in the first and in the second language. Additionally, the effect occurred both when the *talker* (either one with L1 Dutch and L2 English or vice versa), produced speech in their L1 or L2 (note though that overall, reaction times were shorter when the talker used their L1). Importantly, monolingual English control listeners did not show the interlingual homophone effect, ruling out that the effects were due to an accidental confound in the stimuli. Finally, Lagrou, Hartsuiker, and Duyck (2013) embedded the homophones in sentences, and found that even a spoken sentence context does not suffice to render lexical access language-selective.

Interestingly, some studies did show a modulation of cross-linguistic effects by a sentence-level variable, namely semantic constraint, although the results are far from unequivocal. For instance, in the sentence "The handsome man in the white suit is the X" it is not so predictable what X is; but in "The best cabin of the ship belongs to the X", it is perfectly predictable that X is "captain" (examples taken from Van Hell & De Groot, 2008). In a sentence reading task in L2, Van Assche, Drieghe, Duyck, Welvaert, & Hartsuiker (2011) found that semantic predictability did not modulate cognate effects on reading measures. In contrast, other eye-tracking studies (Libben and Titone, 2009; Titone, Libben, Mercier, Whitford and Pivneva (2011) did find modulations of semantic contraint, and so

did studies using tasks like lexical decision or translation (Schwartz & Kroll, 2006; Van Hell & De Groot, 2008). However, also in those studies, sentence contraint did not always modulate cross-linguistic effects, and in the eye-tracking studies this modulation was restricted to late measures such as total reading time.

Finally, the only study that suggested some influence of linguistic cues is Elston-Güttler et al. (2005). In this study, the language cues were extrinsic to the target stimuli themselves, in contrast to the previous ones in which the sentence context in which the stimuli occurred provided the cue. Specifically, the authors manipulated the language of the intertitles in a silent film (i.e., an episode of Louis Feuillade's 1915-1916 film "Les Vampires"). The subjects watched this film while they were being prepared to take part in the EEG experiment and saw a version of it with either L1 German or L2 English intertitles (note that any cues from the actors' lip movements would have cued French, and not the languages of interest here, German and English). After exposure to either the German or English version of the film, the subjects saw sentences followed by target words and conducted a lexical decision on each word. Crucially, in experimental trials the L2 English sentence ended in a word that was a homograph between English and German (e.g., *gift*; German *Gift* means poison), and the target word was related (e.g., the English word *poison*) or unrelated to the German reading of the homograph. Both reaction times (i.e., faster lexical decision times) and event-related potentials (i.e., a reduction of the N400 component) showed that the L1 German reading of the target words was activated. But importantly, these effects were only observed when the subjects had prior seen the German-language

version of the film, and only in the first block. Thus, these findings indicate that extrinsic cues such as the language that is used in a (task-irrelevant) film, can influence the extent to which readers zoom into a language. That the effect did not extend to the second block of the experiment, however, indicates that such an effect is short-lived.

Summarizing, studies that embedded words with cross-linguistic overlap, such as interlingual homographs, interlingual homophones, and cognates in a sentence context found little evidence that the language of the sentence exerted strong constraints on whether lexical selection is language selective or not (with the possible exception of words that are highly predictable in the sentence context). Additionally, although the speech signal is very rich and conveys much information about the speaker, including age, gender, social status, dialect, and native language (e.g., Van Berkum, Van den Brink, Tesink, Kos, & Hagoort, 2008) it seems that cues about the language that is spoken are not enough to rule out the activation of the other-language reading of interlingual homophones (Lagrou et al., 2011; 2013). Note finally that a more in-depth review of this literature is provided in Van Assche, Duyck, and Hartsuiker (2012).

3. Do visual language cues help bilinguals to zoom into the target language?

A number of studies have considered the question of whether visual cues for language affect language activation. These share with the Elston-Güttler et al. (2005) study discussed in the previous section that they are extrinsic to the stimuli that are processed. On the other hand, while Elston-Güttler tried to cue a

language "mode" that was, in principle, irrelevant for the task involving the critical stimuli, visual stimuli can sometimes be more directly linked to language. This is so because people's visual appearances inform us about who they are, allowing us often to infer which language we can expect them to use. Imagine for instance seeing the face of Mr. Lee. Perhaps Mr. Lee is a good friend with whom you often interact (say in English). One possibility then is that seeing Lee's face provides a strong cue about the language you typically use with him. It is also possible that you're introduced to Mr. Lee for the first time. In that case, more general properties of Mr. Lee's appearance might provide language cues even before the conversation has begun – if Mr. Lee looks Asian, you might expect him to use an Asian language such as Chinese, but if he looks Caucasian, his features might lead you to expect him to speak English. Another possibility is that Mr. Lee is in fact the late, Bruce Lee, who was a famous actor in Kung Fu films. In that case, Mr. Lee's appearance might cue you to expect English (if you typically watch the English versions of his films); but if you typically watch Bruce Lee films in Cantonese, Lee's face might cue you to expect that language.

One study that considered the language associated with famous individuals was reported by Hartsuiker and De Clerck (2009). These authors elicited language intrusion errors (e.g., producing Dutch *en* instead of its English translation equivalent *and*; for a different paradigm that induces such intrusions see Gollan, Schotter, Gomez, Murillo, & Rayner, in press). To do so, the authors presented their subjects with triplets of pictures of famous people's faces (e.g., Elvis Presley, Eddie Murphy, and Jennifer Aniston; Figure 1). On each trial, there was an animation, so that for instance two pictures moved in a downward

direction while the third picture stayed put. The subjects, Dutch-English-French trilinguals, described these animations either in their first language Dutch or in one of their non-native languages (English or French), for instance with (2) and (3).

*(2) Elvis **and** Eddie Murphy move down, but Jennifer Aniston stays put*

*(3) Tom Boonen **and** Kim Clijsters move down, but King Albert stays put*

INSERT FIGURE 1 ABOUT HERE

Importantly, this paradigm creates a context around the conjunction *and* (in boldface in 2-3) in which the language of use and the language associated with the famous people "sandwiching" *and* is either congruent (2; Elvis and Eddie Murphy are associated with English) or *incongruent* (3; at least for the subjects tested, students in the Flemish region of Belgium, the famous cyclist Tom Boonen and tennis player Kim Clijsters are associated with Dutch). Of course, in Dutch-language versions of these sentences congruency is flipped: the individuals mentioned in (2) are then language-incongruent and in (3) language-congruent.

In their first experiment, Hartsuiker and De Clerck (2009) observed that indeed, language intrusions happened from time to time in this paradigm. Such intrusions were more likely in the incongruent than the congruent condition and more likely when the target language was L2 than when it was L1. Because the authors were concerned that English *and* and Dutch *en* are near-homophones,

making any language-intrusion errors difficult to detect in speech, this experiment used written production. But follow-up experiments with phonologically distinct conjunctions (i.e., Dutch *en* and French *et*) generalized the language congruency effect to spoken language as well as written language. Further experiments also generalized the findings to a different type of connective (both "or" and "and"). Interestingly, and in agreement with Gollan et al.'s (in press) findings, a final experiment showed that language intrusions occurred more often with function words (e.g., *and*) than with content words (e.g., *cat).* One possible reason is that function words are more likely to be ignored by the processing systems that check our speech and writing for accuracy (i.e., our self-monitoring systems), just as function words are more likely to be skipped than content words during reading (Rayner, 1998). At this point though, this account remains speculative.

The results of this study are suggestive that language-information associated with famous people's faces is activated during language production. However, the study does not allow us to determine whether it is specifically the famous person's face that activates the language, or whether (alternatively or additionally) there is an effect of the person's *name*. That is, one might argue that someone using a name in a different language from the language of the sentence is making a temporary switch to a different language, and it may be difficult to switch back. It is important to note that Gollan et al.'s (in press) experiment is consistent with an effect of names only: that experiment involved reading a text (hence without any visual cues) and also observed language intrusions in mixed language contexts.

It is possible that the case of famous people is a special one. While we may have seen and heard these people in the media, we usually do not interact with them. But in real life, we often have to select a language to use with new people we meet and interact with. Is a short conversation enough to link a language to a person? Martin, Vandenbulcke, Navarra, Schoonbaert, Hartsuiker, and Costa (2011) asked this question in two experiments, one in Barcelona, Spain using Catalan-Spanish speakers and one in Ghent, Belgium using Dutch-French speakers. The experiments had two phases. First, participants saw a video of a person under the pretense that this was a session on Skype (in fact, the video was prerecorded). The person introduced his or herself and spoke about their daily lives and interests, and then invited the subject to likewise make a short speech. The person on the video consistently spoke either Catalan or Spanish (Dutch or French in the second experiment). Next, another person introduced themselves in the same way, so that each participant was acquainted with several Catalan and several Spanish speakers.

The second phase of the experiment was a language production task – given a cue word, provided by the same people previously seen on the video, the participants produced an association as quickly as they could; production latency was measured. Each of the people from the video produced two words in the language they used on the video and two words from the other language. The subjects were instructed to use the language of the cue word. If a short "chat" on a computer-based communication system is enough to create an association between the person and the language they used, one would expect faster production latencies when the language of the cue is congruent with the

language of the video. In contrast, there was no overall effect. A closer look at the data, however, suggested a language congruency effect for the initial trials of the production task. This suggests that while exposure to someone speaking a particular language may lead to an association between that person and the language, such an association can be undone after only a few trials in which that person used another language (i.e., only a few incongruent trials are enough to undo the association that was established before).

In addition to effects of the language associated with specific individuals, it is also possible that, specific *groups* of individuals are associated with particular languages. Two recent studies have recently explored whether the facial features (Asian or Caucasian) of a person whose picture was displayed along with the linguistic stimulus affected processing of a language that was congruent (Chinese) or incongruent (English) with those features. Li, Yang, Scherf, and Li (in press) conducted an fMRI experiment in which Chinese-English bilinguals and a control group of English monolinguals named line drawings of objects. The bilinguals named these stimuli using Chinese or English, depending on a color cue (i.e., a red or blue frame around the picture cued the language to use). Importantly, experimental stimulus displays also showed the face of either Asian or Caucasian persons as well as part of their body; they appeared to be holding the frame with one hand and pointing to it with their other hand (Figure 2). In a control condition, the frame appeared by itself. Thus for the bilinguals, there were 6 naming conditions (3 face conditions and 2 language conditions); for the monolingual group there were only the three face conditions.

INSERT FIGURE 2 ABOUT HERE

Li et al. (in press) observed an interaction between naming condition and language in the analysis of naming latencies, so that participants were fastest when they responded in their L1 (Chinese) and an Asian face accompanied the stimulus. The fMRI data revealed a number of brain areas, mostly in the frontal and temporal lobes, that were more active when using L1 than L2. There were also several areas of the brain more active when seeing faces compared to the no-face control condition, mostly in occipital areas and the fusiform gyrus. Importantly, an analysis of language-congruent vs. –incongruent conditions showed more activation in more areas of the brain (mostly frontal and temporal areas) for the congruent conditions; the effects were strongest when the language was Chinese and the face was Asian.

Finally, the authors also conducted an analysis of four regions of interest, based on earlier brain imaging work that had mapped language control networks (e.g., Abutalebi, 2008). Interestingly, two of these regions (the Medial Frontal Gyrus and Anterior Cingulate Cortex) showed language x face interactions, so that these areas were activated above the no-face baseline in the congruent conditions, and deactivated in the incongruent conditions. The authors interpreted the higher and more extended activation in the congruent condition as a facilitation effect, in line with the facilitation observed on naming latencies. This is somewhat counterintuitive, as one might expect the language control network to be more highly active when there is a conflict between different cues (i.e., face and color cue). The authors suggested that this facilitation reflects a successful process of integration of multiple cues.

A priming effect from Chinese vs. Caucasian faces was also observed by Zhang, Morris, Cheng, and Yap (2013). These authors had subjects (Chinese-English bilinguals) engage in a computer-mediated spoken conversation in English with someone introduced as Michael Lee. At the same time, they saw a picture of either a Caucasian or a Chinese male face, and the authors collected two measures of production fluency, namely subjective fluency ratings and a count of fluently spoken words per minute. The authors reported numerically small but statistically significant cultural priming effects: when the subjects saw the language-incongruent (Chinese) face, their speech production was less fluent on both the subjective and objective fluency measures.

Interestingly, Zhang et al.'s (2013) next experiment replaced the faces with pictures of cultural icons, such as the Great Wall in China or Mount Rushmore in the United States (Figure 3). The Chinese and American cultural icons were equally familiar to the subjects. In a first phase of the task, the subjects (23 Chinese-English bilinguals) described the cultural icons; next they described a set of culture-neutral images. The fascinating result was that in both tasks and for both fluency measures, the participants who saw Chinese cultural icons were significantly less fluent than the ones who saw American icons. Further experiments extended these cultural priming effects to a different domain, namely that of the lexicon. Subjects that were primed with Chinese icons could more quickly identify a literal Chinese translation (e.g., HAPPY NUT for pistachio) and were more likely to produce such a literal translation as an undesired intrusion. Thus, these studies suggest that not only people's faces, but also visual icons of people's cultures can be associated with a language. Some

care needs to be taken in interpreting these results however, given that some of the experiments were rather underpowered, the results seemed to be small (unfortunately, no measure of effect size was reported), and some of the measures (fluency ratings) were subjective. Additionally, the study did not take into account the proficiency of the participants and the objective measure of fluency (i.e., speech rate) may not be the best indicator of second language proficiency (see Yang and Yang, 2013, for a discussion of several concerns with the study and Morris and Zhang, 2013, for a response to these concerns; also see Kroll & McClain, 2013 for further discussion of this study).

INSERT FIGURE 3 ABOUT HERE

In agreement with Zhang et al. (2013), Jared, Pei Jun Poh, and Paivio (2013) found cultural effects on picture naming times in a well-controlled study. These authors presented Chinese-English bilinguals in Canada with objects that either had a Chinese appearance (e.g., a typical Chinese mailbox or cabbage) or a Canadian appearance (a typical Canadian mailbox or cabbage; Figure 4). The subjects named these objects faster if the language used was congruent with the language of the culture to which the picture belonged than in the incongruent conditions. Importantly, this effect occurred both when the subjects named in their L1 (Chinese) and in their L2 (English).

INSERT FIGURE 4 ABOUT HERE

Summarizing, studies in which participants viewed the faces of famous people, of ordinary people they had just interacted with, or with unknown people with either Caucasian or Asian features, converge in showing language

congruency effects. Interestingly, similar effects are also shown with cultural icons as well as with everyday objects that happen to look somewhat different in different cultures.

4. Discussion

The work reviewed here demonstrates compellingly that access to lexical representations, both in the written and spoken modality, both in comprehension or production, is language non-selective. Furthermore, studies that have considered lexical access in a linguistic (e.g., sentence) context, have found that the strong language cues provided by such a context is not enough to restrict lexical selection to only the target language; in particular cognate effects, and effects of interlingual homography or homophony still occur in such contexts. This is especially striking in the case of speech, as this signal is very rich and provides many cues about the talker's identity (age, gender, social status; Van Berkum et al., 2008) and crucially about the language the talker is using. For instance, while the name "Bob" may be spelled the same in English and Spanish, the voice onset time of the initial /b/ will differ according to whether one speaks English or Spanish.

At least at first glance, the studies that so far have considered effects of language-extrinsic cues, such as the visual appearance of people and objects, converge on a rather different conclusion (Hartsuiker & Declerck, 2009; Jared et al., 2013; Li et al., 2013; Martin et al., 2011; Zhang et al., 2013). Specifically, these studies found that the language associated with visual representations affected aspects of language production, including naming latencies, language intrusion speech errors, fluency, and the network of brain areas activated during picture

naming. It is possible perhaps that the visual environment provides cues that more strongly affect representations in the target and non-target language than do linguistic cues. However, before accepting that conclusion it is important to note two major differences between the literatures that considered linguistic and visual cues. First, the literature that considered linguistic cues typically used interlingual effects such as cognate status and homography. No study so far has tested whether such interlingual effects are modulated by visual cues. Is it enough, for instance, to see the face of Elvis Presley to reduce the interference that Dutch-English bilinguals experience when processing Dutch-English homographs? It is crucial to test this, because all comparisons between linguistic and visual cues are now confounded by differences in the measures that the cues are hypothesized to affect.

Second, the literature on visual cues has exclusively focused on measures of language production, while most studies looking at sentence cues have looked at language comprehension. An interesting possibility is that the process of language production is more sensitive to language cues than language comprehension. Such a scenario might be the result of the different demands that language processing in each modality places on the language processing system: the comprehender's immediate goal is to understand the sentence they are presented with. For many practical purposes, it may not be relevant in which language that sentence is (provided of course it is a language the comprehender knows well enough). Thus, for the comprehender a cue for language may not be particularly relevant. This is different for the speaker who *must* make a selection for language for each and every word they produce, which means that it would

be very useful to exploit all language cues that are available to help this selection, including visual cues.

A further issue concerns the mechanisms that underlie the visual language cueing effects we have seen here. One question in particular is whether more fluent production in the context of a language-congruent cue than a language–incongruent cue, results from facilitation (i.e., priming of the correct language representations) or interference (priming the incorrect language representations would hinder the selection of correct language representations). Indeed, Zhang et al. (2013) argue for an interference account, based on the conclusions of one experiment (Experiment 3) in which there were Chinese and American primes (cultural icons) but also matched control conditions. The target language was English. While there was a significant interference in the Chinese prime condition relative to its control condition (i.e., on the latency to recognize literal translations), there was no facilitation in the American primes condition (i.e., on the latency to recognize object names) relative to its control condition. However, Zhang et al.'s experiment constitutes very limited evidence for an interference account. As pointed out by Yang and Yang (2013), the literal translation recognition task is far from ideal, as the task trivially requires access to the native language. Furthermore, the conditions that test for interference (involving literal translations) are different from those testing for facilitation (involving object names). Finally, Zhang et al. only tested production in English, with subjects immersed in a context where English was presumably the default (a university in the United States). It is possible that in such a context, exposure to Chinese cultural icons is rather unexpected and draws attention, while

exposure to American cultural icons is much more expected. Thus, the addition of a Chinese-language condition woud have aided a more reliable assessment of the interference hypothesis.

Fortunately, the picture naming fMRI study reported by Li et al. (in press) contained both a Chinese and an English language condition and had both Chinese and English primes (i.e., faces) and a control condition. The pattern of naming latencies clearly indicated that the effect of face cues is facilitatory in nature. Additionally, the pattern of activation in the brain (i.e., more extensive activation in language control areas in the congruent than incongruent conditions) also suggested a facilitation effect, although it is also possible of course that there is both facilitation *and* interference.

A further issue is how general effects of visual language cues are. For instance, Yang and Yang (2013) suggest that Zhang et al.'s (2013) visual language cueing effects might be restricted to subjects with very low proficiency, based on observations in the literature that grammatical and lexical intrusions of the native language occur more frequently in relatively lower proficiency bilinguals than higher proficiency bilinguals (Poulisse & Bongaerts, 1994). Kroll and McClain (2013) similarly suggest that such effects might well be modulated by factors such as age of second language acquisition and whether the bilingual is immersed in their second language or not (also see above). Future research will have to establish what the influence of such factors is, although it is important to note that some of the studies we discussed so far used extremely high-proficient bilinguals (the Spanish-Catalan bilinguals tested by Martin et al.) that typically acquired their second language early, as well as late bilinguals that were

immersed in an L1-dominant environment (Hartsuiker & Declerck, 2009). Of course, the same question can be asked for effects of linguistic cues too, and as pointed out by Van Assche et al. (2012), those studies also differ in for instance participant characteristics.

Most of the studies discussed so far consider the findings in light of language control, assuming that valid language cues – be they linguistic or visual in nature - can help suppress the irrelevant language. However, the question of which representations in bilingual memory are affected by language cues has so far not been answered conclusively. One possibility is that there are language nodes (e.g., Green, 1986; Van Heuven, Dijkstra, & Grainger, 1998) and that extrinsic language cues directly affect the activation of such nodes. On such an account, if I decide to speak Dutch, but I see a picture of (say) Elvis, this visual cue might activate my language node for L2 English, which would then promote the possibility of interference from English. In Dijkstra and Van Heuven's (2002) BIA+ model, the mechanisms of top-down language suppression by language nodes has been replaced by a so-called task schema, that is sensitive to situational demands (e.g., the specific task the subject has to perform) without directly affecting the activation of representations in the lexicon. A further possibility is then that visual language cues directly affect task schemata, leading for instance to interference when the incorrect task schema is cued.

However, Jared et al. (2013) propose a fascinating alternative account, at least for their findings of a language-congruency effect on the naming of objects with culturally-specific looks, such as a Chinese vs. a Canadian cabbage. Rather than interpreting this effect in terms of language control mechanisms, they frame

the finding in terms of a bilingual dual coding model, which assumes a store of visual conceptual representations (Paivio & Lambert, 1981). Importantly, Jared et al. assume that within such a model, there would be asymmetric connections from the visual representations to lexical items in the two languages. Thus, a visual representation for a Canadian cabbage would be more strongly connected to the English word "cabbage" than to its Chinese translation equivalent, while the reverse would be true for a visual representation of a Chinese cabbage. Therefore, naming the Canadian cabbage in English would be easier than naming it in Chinese, with the reverse holding for the Chinese cabbage. In such an account, at least some language congruency effects can be seen as a direct consequences of the structure of speakers' semantic memory, rather than of language control mechanisms.

In conclusion, bilingual language production is affected by visual cues in the environment that point to one of the two (or more) languages the speaker knows. This is in interesting contrast to the effect of linguistic cues (e.g., sentence context) that has more often been studied in the language comprehension literature. While it may be a while before we understand better when and why such visual language cueing effects occur, the results obtained so far clearly indicate that visual information can directly affect language processing in bilinguals.

Acknowledgment

Hans Rutger Bosker is thanked for very insightful comments on an earlier version of this paper.

References

Abutalebi, J. (2008). Neural aspects of second language representation and language control. *Acta Psychologica,* 128, 466-478.

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

De Bruijn, E. R. A., Dijkstra, T., Chwilla, D J., & Schriefers, H. J. (2001). Language context effects on interlingual homograph recognition: evidence from event-related potentials and response times in semantic priming. *Bilingualism: Language and Cognition, 4, 155-168.*

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

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

Duyck, W., Van Assche, E., Drieghe, D., and Hartsuiker, R. J. (2007). Visual word recognition by bilinguals in a sentence context: evidence for nonselective lexical access. *Journal of Experimental Psychology: Learning Memory and Cognition,* 33, 663–679.

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

Gollan, T. H., Schotter, E. R., Gomez, J., Murillo, M., & Rayner, K. (in press). Multiple levels of bilingual language control: Evidence from language intrusions in reading aloud. *Psychological Science.*

Green, D. W. (1986). Control, activation, and resource: a framework for the control of speech in bilinguals. *Brain and Language,* 27, 210-223.

Hartsuiker, R. J. & De Clerck, M. (2009, September). Albert Costa y Julio Iglesias move up but Fidel Castro stays put: Language attraction in bilingual language production. Paper presented at the AMLaP 2009 conference, Barcelona, Spain.

Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica,* 127, 277-288.

Jared, D., Pei Jun Poh, R., & Pavio, A. (2013). L1 and L2 picture naming in Mandarin-English bilinguals: a test of bilingual dual coding theory. *Bilingualism: Language and Cognition, 16, 383-396.*

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of Bilingualism for language processing and cognitive control. *Journal of Cognitive Psychology.*

Kroll, J. F., & McClain, R. (2013). What bilinguals tell us about culture, cognition, and language. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 11219-11220.

Lagrou, E., Hartsuiker, R. J., and Duyck, W. (2011). Knowledge of a second language influences auditory word recognition in the native language. *Journal of Experimental Psychology: Learning Memory and Cognition,* 37, 952–965.

Lagrou, E., Hartsuiker, R. J., and Duyck, W. (2013). The influence of sentence context and accented speech on lexical access in second-language auditory word recognition. *Bilingualism: Language and Cognition*, 16, 508-517.

Lemhöfer, K., Dijkstra, T., and Michel, M. C. (2004). Three languages, one echo: cognate effects in trilingual word recognition. *Language and Cognitive Processes,* 19, 585–611.

Li, Y., Yang, J., Scherf,K. S., and Li, P. (in press). Two faces, two languages: An fMRI study of bilingual picture naming. *Brain and Language.*

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

Martin, C., VandenBulcke, C., Navarra, E., Schoonbaert, S., Hartsuiker, R.J., & Costa, A. (2011; September). Please, Catalan or Spanish, but not both! Are bilinguals fully in control of their language selection during word production? Paper presented at the Workshop on Bilingualism:

Neurolinguistic and Psycholinguistic Perspectives, Aix-en-Provence, France.

Morris, M. W., & Zhang, S. (2013). Reply to Yang and Yang: Culturally primed first-language intrusion into second-language processing is associative spillover, not strategy. *Proceedings of the National Academy of Sciences of the United States of America, 110.*

Poulisse, N., and Bongaerts, T. (1994). 1st Language use in 2nd-language production. *Applied Linguistics,* 15, 36–57.

Paivio, A., & Lambert, W. (1981). Dual coding and bilingual memory. *Journal of Verbal Learning and Verbal Behavior, 20, 532-539.*

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

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

Thierry, G., & Wu, Y. J. (2007). Brain potentials reveal unconscious translation during foreign-language comprehension. *Proceedings of the National Academy of Sciences of the United States of America, 104, 12530-12535.*

Titone, D., Libben, M., Mercier, J., Whitford, V., and 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., Duyck, W., Hartsuiker, R. J., and Diependaele, K. (2009). *Does bilingualism change native-language reading? Cognate effects in a sentence context*. Psychological Science, 20, 923–927.

Van Assche, E., Drieghe, D., Duyck, W., Welvaert, M., and 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. (2012). Bilingual word recognition in a sentence context. *Frontiers in Psychology*, 3:174.

Van Berkum, J. J. A., Van den Brink, D., Tesink, C. M. J. Y., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience, 20, 580-591.*

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

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

Yang, S., & Yang, H. (2013). Does bilingual fluency moderate the disruption effect of cultural cues on second-language processing? *Proceedings of the National Academy of Sciences of the United States of America, 110.*

Zhang, S., Morris, M. W., Cheng, C-Y., & Yap, A. J. (2013). Heritage-culture images disrupt immigrants' second-language processing through triggering first-language interference. *Proceedings of the National Academy of Sciences of the United States of America, 110, 11272-11277.*

Figure captions

1. Example of the displays used in Hartsuiker and Declerck (2009). In the actual experiment, two of the pictures would move, for instance the left and middle picture could move downwards.

2. Example of the stimuli used in Li et al. (in press). Permission not yet asked to Elsevier

3. Example of the stimuli used in Zhang et al. (2013). Permission not yet asked to PNAS; but these two were listed in Zhang et al. as belonging to the public domain.

4. Example of the stimuli used in Jared et al. (2013). Permission not yet asked to Cambridge University Press.

Figure 1.

Figure 2

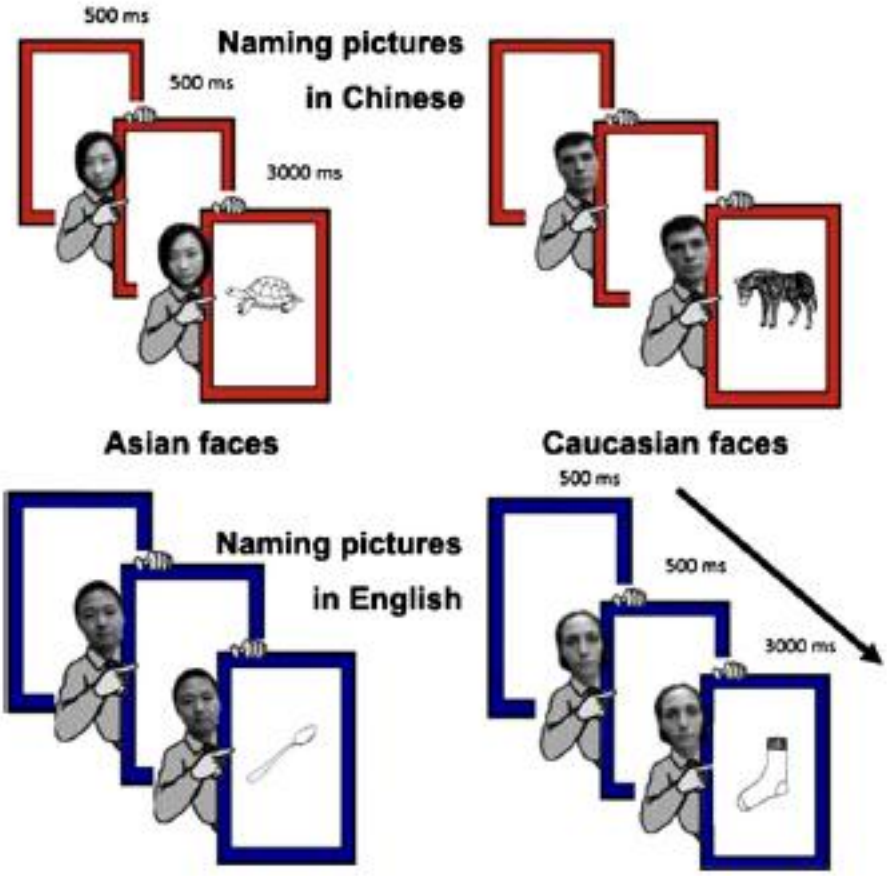
500 ms
Naming pictures
in Chinese
500 ms
3000 ms
Asian faces
Caucasian faces
Naming pictures
in English
500 ms
500 ms
3000 ms

Figure 3

Figure 4