Chapter 7: Lexical selection and competition in bilinguals

Mikel Santesteban\textsuperscript{1} & John W. Schwieter\textsuperscript{2,3}

\textsuperscript{1} University of the Basque Country, Spain  
\textsuperscript{2} Wilfrid Laurier University, Canada  
\textsuperscript{3} University of Greenwich, England

This chapter critically reviews the bilingual ambiguity literature with an emphasis on speech production and cross-language cognates. The chapter also looks at issues related to language competition, language facilitation, the language- and task-switching experimental paradigm, and the extent to which language control is subsidiary of domain-general executive functions. Finally, we discuss ongoing dialogues on the bilingual advantage debate.

Key words: bilingualism, cognate words, language production, language control, bilingual advantage.

1. Introduction

Which single word would you use to describe a long piece of furniture for seating that typically has a back and an armrest? To respond to this question, you would first need to access the conceptual representation corresponding to the description. Next, you will need to access its lexical representation, which would also lead to the activation of its phonological form. Finally you would probably respond “sofa”. Or would you have produced “couch” instead? The main models of lexical selection assume this type of staged lexical access process (Caramazza, 1997; Dell, 1986; Levelt, 2001; Levelt, Roelofs, & Meyer, 1999). Interestingly, there is compelling evidence suggesting that during lexical access, along the process of selecting the target word, other words that might be semantically and/or phonologically related to the target word get also activated. And all activated words might compete during lexical selection. However, since lexical selection is mainly conceptually guided, the target word would be activated more than any other words, and would be finally selected. In the case of synonymy cases as the one mentioned above, factors such as the frequency with which the speaker uses each synonym word would determine the selection process and the speaker’s preference to produce “sofa” or “couch” (Levelt, 2001). But in monolingual speech production, synonymy cases are not so frequent and descriptions as the one above usually lead to unique possible responses. In contrast, in the case of bilingual speakers that have two similarly adequate words for almost every concept, the lexical selection process might be more complex.

2. Lexical selection in bilingual speech production

One of the main research questions in bilingual lexical access is whether the two languages of bilinguals get activated, and if so, what lexical selection mechanisms do
bilinguals use to avoid the competition of the lexical items of the non-response language. As we will review next, there is compelling evidence suggesting that the two languages of bilinguals are activated during lexical access. However, there is still hot debate about the nature of the mechanisms controlling lexical selection in bilingual speech production. Do the co-activated words of the other language compete for selection? How do bilinguals select words in the language they aim to use and avoid the activation of the words in the non-response language? Next we briefly summarize the main research that aimed to answer these questions.

### 2.1. Co-activation of both languages

Most bilingual lexical access models assume that during the course of lexicalization in one language, the semantic system activates the lexical nodes of the two languages of bilinguals (e.g., Colomé, 2001; Costa, Caramazza, & Sebastián-Gallés, 2000; Costa, Colomé, Gómez & Sebastián-Gallés, 2003; Costa & Santesteban, 2004; Gollan & Kroll, 2001; Hermans, Bongaerts, De Bot & Schreuder, 1998). Evidence for the co-activation of the two languages comes from various experimental paradigms. For instance, using a phoneme-monitoring task, Colomé (2001) demonstrated that the activation of the non-target lexical representations spreads to their phonological representations. In this task, participants have to determine whether a given phoneme is present in the name of a target picture. Her main findings showed that, upon the presentation of the picture of a table ("taula" in Catalan and "mesa" in Spanish), Catalan-Spanish highly-proficient bilinguals were faster determining that the phoneme “t” was present in the Catalan name “taula” than determining that the phoneme “m” was not. More importantly, when negative responses where required, participants needed more time to assess the absence of the phoneme “m” (not present in “taula” but present in its Spanish translation: “mesa”) than of the phoneme “f” (not present in either “taula” or “mesa”). These results were interpreted as revealing the simultaneous activation of bilinguals’ two languages up to sublexical phonological levels.

Further evidence comes from the picture-word interference paradigm, in which participants are asked to name a picture while ignoring an auditorily- or visually-presented distractor word. In this paradigm, two main effects have been used to inform about lexical access processes: The semantic interference effect stands for the slower naming latencies observed when the distractor word is semantically related (e.g., cat), as compared to when is unrelated (e.g., chair), to the picture to be named (e.g., dog). The slow-down is assumed to occur due lexical competition between the activated lexical items during lexical selection. The phonological facilitation effect stands for the faster naming latencies observed when the distractor is phonologically related (e.g., doll) to the name of the picture (dog). The effect is assumed to occur after lexical selection, during the retrieval of the word’s phonological form, and occurs because the distractor word enhances the activation of the phonemes shared with the name of the picture. Interestingly, Hermans et al. (1998) asked Dutch-English bilinguals to name pictures in their L2-English (e.g., “mountain”) while presenting English distractor words that were either unrelated to the name of the picture (e.g., “present”) or phonologically related to the picture name’s Dutch translation (e.g., “bench”, which is phonologically
related to “berg”, mountain in Dutch). These latter type of distracters slowed-down participants' responses as compared to unrelated distracters, revealing the so-called phono-translation effect (Hermans et al., 1998; see also, Costa, et al., 2003; Costa, Albareda, & Santesteban, 2008). This effect is interpreted to occur because the distractor “bench” further activates the already co-activated picture translation’s name (“berg”). As a consequence, and assuming that the words of the non-target language compete during lexical selection, the Dutch lexical node “berg” would be a more powerful competitor to the English lexical node “mountain” when the picture is presented with the distractor “bench” than when presented with the unrelated distractor word “present”. Hence, these results suggest not only that there is co-activation of the two languages of bilinguals, but also that these languages compete for selection. Next we discuss further evidence in support and against cross-linguistic competition.

2.2. Cross-language competition and cognates

Despite the agreement regarding the existence of simultaneous activation of the two languages, it is still debated whether the two languages compete for selection. The results of Hermans et al (1998) suggest that they do. However, Costa and colleagues provide the most compelling evidence suggesting that there might not be such cross-linguistic competition. In a series of picture-word interference experiments, they showed that bilingual speakers name pictures faster when they are presented with a distractor word that corresponds to the target’s translation than when it is an unrelated word (Costa and Caramazza, 1999; Costa, Miozzo, & Caramazza, 1999; see also Hermans, 2004). For instance, English-Spanish bilinguals are faster naming the picture of a dog in Spanish (“perro”) when the distractor is “dog” than when it is “chair” (both presented in English). This so-called cross-language identity effect is at first sight paradoxical given that if there were to be competition across languages, the lexical node corresponding to the target’s translation should be the most powerful competitor. This is because, if within-language semantic interference effects reflect lexical competition of semantically related words, translation words should be expected to create the maximal competition since their semantic overlap with the target word is also maximal. Hence, translation word distracters should slow down picture naming latencies rather than facilitate them. Researchers suggested that facilitation is observed because there is no cross-linguistic competition, and since the target name is activated through both the picture and the distractor word’s translation, its selection occurs faster.

Based on these findings, Costa and colleagues proposed a language-specific selection model according to which, although the bilinguals’ two languages get activated, the activated non-target language’s lexical representations do not compete for selection. In other words, according with this account, bilingual speakers’ lexical selection mechanism is sensitive only to the activation of the lexical items of the to-be-produced language and ignores the activation of the non-target language’s words without the need to suppress their activation. Alternatively, the Inhibitory Control (IC) model proposed by Green (1998), also referred as language-non-specific selection model, assumes that the two languages of bilinguals compete for selection, so that bilinguals use inhibitory mechanisms to suppress the activation of the non-target language.
Regarding whether there is cross-language competition in bilingual lexical selection, there is one feature of translation words that might suggest there is not: the cognate status of words. Cognates are those words that are phonologically similar in the bilingual’s two languages (e.g., “guitarra” and “guitar” for Spanish-English bilinguals) while non-cognates are phonologically dissimilar (e.g., “mesa” and “table”). The cognate status of words has pervasive effects in the performance of bilinguals. For instance, in picture naming, Costa, et al. (2000) showed that bilingual speakers are faster naming pictures with cognate names than pictures with non-cognate names (see also, Christoffels, Firk, & Schiller, 2007; Hoshino & Kroll, 2008; Strijkers, Costa, & Thierry, 2010). Gollan and Acenas (2004) also showed that bilinguals are less likely to fall in tip-of-the-tongue states while naming cognates than non-cognates, suggesting that cognates are more resistant to momentary malfunctioning of the lexical retrieval mechanism. Additionally, cognate words are also faster to learn and more resistant to forgetting (e.g., de Groot & Keijzer, 2000; Leacox, Wood, Sunderman, & Schatschneider, 2016), more sensitive to cross-linguistic priming (de Groot & Nas, 1991; van Hell & de Groot, 1998), and more resistant to anomic states in aphasic patients and to deterioration in patients with Alzheimer’s disease (Costa et al., 2012; Roberts & Deslauriers, 1999). Cognate words seem to elicit a more similar brain activity between languages (De Blesser, Dupont, Postler, et al., 2003) and show more language transfer in rehabilitation of aphasic patients (Kohnert, 2004) and in intervention of children with specific language impairment (Kambanaros, Michaelides, & Grohmann, in press).

These cognate facilitation effects have been interpreted as evidence for the co-activation of bilingual’s two lexical elements up to their phonological representations as well as for language specific selection models (Costa et al., 2000; Costa, Santesteban, & Cañó, 2005). According to this phonological account of the origin of cognate effects, facilitation effects occur because the phonological form of cognates gets activated from two sources, the target word and its co-activated phonologically related translation, and this activation feeds back to the lexical level, facilitating lexical selection. Alternatively, Kroll, Bobb, and Wodniecka (2006) suggested that bilingual speakers might rely on inhibitory control mechanisms for language selection, but that lexical selection (and competition) might occur at different levels of representation depending on different factors such as the degree of activation of the non-target language or the bilinguals’ proficiency level. Thus, cognate effects would suggest that lexical selection and competition might occur at the phonological rather than the lexical level of representation (probably due to the enhanced level of co-activation of the target’s translation word). See section 5 for a more thorough review of these type of language control mechanisms.

However, an alternative hypothesis about the origin of the cognate facilitation effects suggests that they originate because cognates share more semantic features (i.e., have a larger semantic overlap) than non-cognates (van Hell & de Groot, 1998). Other authors have also claimed that cognates might share, to some extent, their lexical (morphological) representation (Kirsner, Lalor, and Hird (1993), although this proposal is
more problematic, and would be limited to identical cognates (e.g., “piano” in both Spanish and English). But note that even these identical morphological forms usually lead to different phonological realizations (see Costa, Santesteban, & Cañó, 2005, for more detailed discussion on the origin of cognate effects).

Recent electrophysiological evidence in picture naming showed that the cognate status of words affects lexical factors such as lexical frequency, and elicits early ERP effects that would be difficult to reconcile with a purely phonological origin of the cognate effects (Christoffels, et al., 2007: effects between 275-375 ms; Strijkers et al., 2010: a P2 around 180 ms). Based on a meta-analysis of behavioral and ERP data, Indefrey and Levelt (2004) estimate that access and retrieval of lexical information in picture naming occurs at a time window between 175 and 330 ms after picture onset, while access and retrieval of postlexical (phonological) information would occur from 330 ms onwards. However, Strijkers and Costa (2016) recently proposed a neural assembly model of language production according to which words are not represented and accessed in a staged hierarchical manner, but as word assemblies that ignite as a whole, activating semantic, lexical and phonological information simultaneously. Interestingly, Strijkers (in press) has recently proposed a bilingual version of this assembly-based model that accounts parsimoniously with the early ERP effects elicited by cognate effects. According to this account, in the brains of bilingual speakers, the neural assemblies of non-cognate translation words would be overlapping at the sites at which lexico-semantic representations are stored, but not much overlapping will occur at sites at which lexico-phonological representations are stored. However, the representations of cognate words would be highly overlapping at both lexico-semantic and lexico-phonological sites, and the higher the cognate words' phonological overlap, the higher the overlap in L1 and L2 assemblies (i.e., with larger overlapping at the brain sites where the phonological information of words is stored for identical than non-identical cognates).

In sum, although the origin of cognate effects and the way they are represented in the brain is still debated, there is wide agreement that cognate effects reveal the co-activation of the two-languages of bilinguals.

3. Language competition and facilitation

As mentioned above, the presence of cognate facilitation effects in the naming performance of bilingual speakers suggests that the two languages of bilinguals might not compete for selection. However, other bilingual effects suggest that the activation of the non-target language affects their performance, so that some sort of control mechanism needs to regulate the cross-linguistic activation and/or competition. For instance, one of such effects is the language mixing cost. These costs are calculated by comparing the bilingual’s L1 and L2 picture naming performance in a context in which they only use one of their languages (L1 or L2 blocked language context) versus a context in which they are required to use both languages interchangeably (language switching context). The difference between switching and blocked language conditions is referred to as the mixing cost, and the cost is calculated by comparing the response
times to language repeat trials in the language switching task (e.g., trials on which participants’ respond in the same language as in the previous trial, non-switch trials) with responses in blocked naming contexts. The rationale is that the mixing costs resulting from the comparisons between these two naming contexts would be reflecting the bilinguals’ efforts to deal with cross-linguistic competition.

The main pattern of results observed so far is diverse: some studies report mixing costs in both the L1 and the L2 (e.g. Prior & Gollan, 2011), others only in L1 but not in L2 (e.g. Christoffels, et al., 2007); others even obtained mixing costs in L1 but a “mixing benefit” in the L2 (e.g., faster responses in the mixed than blocked language contexts: Gollan & Ferreira, 2009; Mosca & Clahsen, 2016) However, the main pattern of results suggests that mixing costs are larger in the dominant L1 than in the non-dominant L2, and these findings have been interpreted as evidence that bilinguals make use of global inhibitory mechanisms to prevent the competition of the non-target language: the larger L1 mixing costs would be reflecting the need to inhibit to a larger extent the dominant-L1 while speaking in the non-dominant L2 (for further discussion, see Kroll, Bobb, Misra & Guo, 2008; Philipp & Koch, 2009). Interestingly, Christoffels et al. (2007) used the cognate facilitation effect as an index of the level of co-activation of translation words in blocked and mixed language contexts. They observed both that mixing costs only occurred in L1 but not in L2 and that naming in mixed language conditions enhanced the magnitude of cognate effects in L1 but not in L2. The researchers interpreted these findings as evidence that the activation of the dominant L1 is lowered by means of global inhibitory processes. Consequently, the influence of the co-activated L2 words would be enhanced, leading to larger cognate effects.

Similar arguments have been made in regard to the L2-immersion effects shown in the bilingual speakers’ L1 naming performance. Linck, Kroll and Sunderman (2009) showed that L2 learners suffer certain difficulties in L1 lexical retrieval after intensive L2 practice in a L2-immersion context (e.g., living in an L2 speaking country). More specifically, in a verbal fluency task, two groups of English learners of Spanish matched in length of L2 study were asked to produce as many words of a semantic category as they could. The L2 immersed learners that lived in Spain for one semester produced more words in Spanish than the classroom learners that studied Spanish in the EEUU. More importantly, when they were tested in their L1-English, the L2 immersed learners produced fewer words than the classroom learners. These results were interpreted as evidence that in the L2 immersion context, the L1 is actively inhibited.

However, Baus, Costa, and Carreiras (2013) suggested that these L2-immersion effects might be rather due to the changes in frequency of use of the L1 during the L2-immersion period. These authors tested the changes in L1-German performance of a group of German learners before and after the L2-immersion period. More specifically, they tested whether L2-immersion effects were modulated by lexical frequency and the cognate status of words, and observed that L2-immersion effects were restricted to words used with low frequency and that were non-cognates: In a picture naming task, German speakers showed slower naming latencies at the end than at the beginning of the immersion period only for pictures whose names were low frequency and non-
cognates. In a semantic fluency task, they revealed a decrease in the percentage of words produced at the end of the L2-immersion period for non-cognate words, but not for cognates. The fact that L2-immersion did not affect the accessibility to the L1 lexicon as a whole suggests that the L1 might not have been actively inhibited. Instead, Baus and colleagues suggested that their findings are more in line with theoretical accounts that consider that frequency of use of words as a key factor determining lexical accessibility. Since the L2-immersion period changed the frequency of use of the L1, it also reduced the availability of those L1 lexical representations that are produced less frequently and are co-activated to a lesser extent due to their non-cognate status. Probably the most representative account of this type is the so-called weaker links hypothesis that we present next.

3.1. Bilingual “disadvantages” and the weaker links hypothesis.

The weaker links hypothesis was proposed by Gollan and collaborators to account for the “disadvantages” observed in the performance of bilingual speakers as compared to monolingual ones during lexical access in their native language. More specifically, in several picture naming and fluency tasks, researchers have observed that, even when speaking in their L1, bilingual speakers are slower naming pictures and produce fewer words than monolinguals (Gollan, Fennema-Notestine, Montoya, & Jernigan, 2007; Gollan, Montoya, Cera, & Sandoval, 2008; Gollan et al., 2011; Ivanova & Costa, 2008; Sandoval, Gollan, Ferreira, & Salmon, 2010; for a review see Runnqvist, Strijkers, Sadat, & Costa, 2011). The weaker links hypothesis assumes that such a bilingual disadvantage occurs because bilinguals have two different words for almost all given concepts, one per language, so that bilinguals would make use of each of these words less frequently than monolingual speakers. Due to the less frequent use, the links between the semantic and phonological representations of words in the two languages of bilinguals are weaker than those in the only language of monolinguals. Thus, considering the logarithmic relationship between lexical frequency and naming speed, frequency effects should have a bigger impact on rarely used words than frequently used words. Being so, the weaker links hypothesis predicts that bilinguals should show larger frequency effects than monolinguals (i.e., a greater disadvantage for low-frequency than for high-frequency words).

This is in fact what Gollan et al. (2008) obtained in an English picture naming task in which the performance of Spanish-English bilinguals with L2-English as their dominant language was compared to that of English monolinguals. This result was also replicated by Ivanova and Costa (2008) with bilinguals speaking in their dominant L1: Spanish-Catalan bilinguals were slower naming pictures in their dominant L1-Spanish than Spanish monolinguals. This bilingual disadvantage effect was modulated by lexical frequency, with effects being larger for low-frequency than high-frequency words (and concurrently, frequency effects were larger for bilinguals than for monolinguals).

Interestingly, the bilingual disadvantage disappears or gets reduced when bilinguals produce words that may be shared across languages, like proper names, or that are phonologically related, such as cognate words. For instance, Gollan and colleagues
showed that bilinguals experience more tip-of-the-tongue states than monolinguals, except when they have to produce proper names or cognates (Gollan & Acenas, 2004; Gollan, Bonanni, & Montoya, 2005). In lexical fluency tasks, although no bilingual disadvantages were found, Blumenfeld, Bobb, & Marian (2016) recently showed that Spanish-English bilinguals speaking in English, their dominant language, retrieve a higher percentage of cognates than monolinguals (see also Sandoval et al., 2010). The level of proficiency in the bilinguals’ non-dominant Spanish also seemed to affect their lexical fluency in English: the higher their proficiency level in Spanish, the higher their preference to produce cognates in English. Additionally, the cognate status of words seemed to affect the bilinguals’ lexical fluency performance in their two languages: the lexical frequency of produced cognate words was similar in the two languages, but in the case of non-cognates, bilinguals’ produced non-cognates with lower frequency values when the task was performed in the non-dominant Spanish than in English. These findings reveal that bilingual lexical access, even when speaking in the dominant language, is affected by the level of proficiency in the non-dominant language, which also enhances the production of cognate words. This is probably because the larger the proficiency level in the non-dominant language the larger its co-activation level during the production of the dominant language. This in turn would also enhance the production of words with larger co-activation levels, such as cognate words, that might also overcome lexical frequency effects.

Certainly, further research needs to explore the extent to which language proficiency changed the way bilinguals access cognate words in their lexicons, and the extent to which the cognate status of words interacts with frequency effects in bilingual lexical access. Next we present a theoretical account that considers the way the cognate status of words might affect the co-activation levels of the two languages of bilinguals and suggests that the production of cognate words might trigger within-sentence code-switching in bilingual communication.

3.2. Code-switching and the cognate triggering hypothesis

Clyne (1967; 2003) analyzed language contact and code-switching corpora of several types of bilingual populations and observed that bilingual speakers seem to be more prone to switch languages in the vicinity of cognate words. Based on this observation, he proposed that cognate words might trigger language switching. Broersma and De Bot (2006) further elaborated Clyne’s proposal, and hypothesized that, since cognate words share most of their phonological form, the production of a cognate word might increase the activation of lexical entries of the other language more than the production of a non-cognate word. Hence, when a bilingual is speaking in language A and produces a word that is a cognate with a lexical entry in language B, the level of activation of language B increases, which might trigger a switch from language A to language B. Researchers proposed that triggering takes place at the lexical (lemma) level of representation. Thus, assuming that the cognate effects arise at the phonological level, the non-target lexicon would be assumed to be activated through feedback activation from the phonological to the lexical level of representation (Goldrick, 2014).
Broersma and collaborators provided corpus analyses evidence supporting this so-called “cognate triggering hypothesis” (Broersma, 2009; Broersma & De Bot, 2006; Broersma, Isurin, Bultena, & De Bot, 2009). For instance, Broersma and De Bot (2006) analyzed the data of a corpus (of 2224 words from one speaker) containing conversations between three Dutch-Moroccan Arabic bilingual speakers. Their results showed that language switching occurred more frequently for words that immediately followed a cognate word or that were part of the same clause as the cognate than words that did not have a cognate word in their vicinity. Broersma, et al. (2009) replicated these findings in their analyses of corpora of interviews to elderly Dutch-English bilinguals from New Zealand (13648 words from four speakers) and Australia (9344 words from two speakers) and a corpus of an interview with one Russian-English bilingual from the United States (2896 words). These findings revealed that cognate words trigger language switching regardless of their grammatical class and regardless of the amount of lexical overlap (i.e., number of cognates) and the typological similarity of the languages spoken by the bilinguals. Additionally, evidence from syntactic priming (i.e., the tendency of speakers to repeat syntactic structures) suggests that cognate words can enhance intra-sentential codeswitching (Kootstra, van Hell, & Dijkstra, 2012). These authors showed that Dutch-English high-proficient bilinguals tended to produce more intra-sentential (Dutch to English) language switches at the same sentence position as in the prime sentences when sentences contained a cognate than a non-cognate word.

In contrast, in two studies, Bultena, Dijkstra, and van Hell (2015a, 2015b) provide negative evidence for the “triggering hypothesis”. They explored whether the presence of a cognate verb facilitates the comprehension and production (shadowing) of intra-sentential language switching. Dutch-English bilinguals were tested with self-pace reading (Bultena et al., 2015a) and sentence shadowing (Bultena et al., 2015b) paradigms on which they were asked to read or shadow (repeat instantaneously) L1-to-L2 or L2-to-L1 code-switched sentences. Results showed that the cognate status of the verb did not modulate the costs of switching languages in any direction (neither in reading times or in shadowing latencies) at the following sentence position.

Finally, the two studies that directly explored the “cognate triggering hypothesis” at the lexical level of representation are not conclusive (Broersma, 2011; Santesteban & Costa, 2016). Both studies used the language switching paradigm (see section 5 for a detailed description of the task). In two experiments, Broersma (2011) tested the language switching performance of medium-high proficient bilinguals of L1-Dutch and L2-English. In cued language switching contexts, participants were faster switching languages after having produced cognate than non-cognate words. When bilinguals were allowed to switch languages at will, they switched more frequently following a cognate than a non-cognate. These findings support the “triggering hypothesis” during lexical access. However, Santesteban and Costa (2016) failed to replicate these findings with low-proficient and high-proficient L1-Spanish and L2-Catalan bilinguals: they showed similar language switching costs before and after naming cognate and
non-cognate words (see also Christoffels et al., 2007; Verhoef, Roelofs, & Chwilla, 2009).

In sum, overall findings of code-switching at the sentential level seem to support the triggering hypothesis proposed by Broersma & De Bot (2006), while language switching data at the lexical level are not conclusive. One of the main claims of the “cognate triggering hypothesis” suggests that the cognate status of words modulates the level of activation of the non-target language of bilinguals, which in turn is assumed to affect the language control mechanisms used by bilinguals for language selection, so that the production of cognates might prompt bilingual speakers to switch languages during conversation. Certainly, future research needs to further test the assessments of the “cognate triggering hypothesis”, and need to explore whether the cognate status of words might affect the way bilinguals control their languages. But we will leave this for future research.

So far we have mainly focused on discussing the way lexical factors such as the cognate status of words might affect bilingual lexical access. In the following sections we will present in some more detail the IC model (Green, 1998) mentioned above, probably the most influential model of bilingual language control accounting for the way bilinguals avoid the interference of their non-target language. We will present the language switching paradigm, a method extensively used to support the claim that inhibitory processes are used in language control. Finally, we will discuss the extent to which language control mechanisms make use of general cognitive control mechanisms.

4. Language control and other cognitive domains

Speech production is inherently guided by top-down control processes—the speech act is initiated by the speaker’s intention to communicate, and speakers monitor their output to ensure alignment with their interlocutors (Levelt, 1989). But as activation spreads through the language system, numerous possibilities arise for conflict between competing representations, both within a language (Andrews, 1997) and, in the case of multilinguals, between languages (Costa, et al., 1999; Hermans, 2000; Hermans, et al., 1998; Kroll, et al., 2006). Thus, at multiple levels within the language system, there is a need for cognitive control mechanisms to support multilingual speech production. Behavioral studies of multilingual language control have separately examined the importance of different executive functions, including inhibitory control (IC; Koch, Gade, Schuch, & Philipp, 2010; Linck, Schwieter, & Sunderman, 2012), working memory (WM; e.g., Christoffels, de Groot, & Waldorp, 2003), and task switching (Prior & Gollan, 2011) (see Festman & Schwieter, 2015 for a review).

Cognitive control has been argued to operate by various mechanisms. On the one hand, global or sustained control guides behavior by keeping active the current goal of the system (Braver, Reynolds, & Donaldson, 2003). Such global control is necessary to ensure the goal-appropriate response is selected, particularly when multiple task-relevant responses are available for response selection. However, even with a clear
goal in place, competition occurs in many forms, such as between distractor and target representations, between competing responses, or between different dimensions of the target representation (see the Dimensional Overlap Model, Kornblum, Hasbroucq, & Osman, 1990). With language tasks, even more opportunities for conflict appear throughout the system at the phonological, orthographic, and morphosyntactic levels (Kroll et al., 2006). Some additional local control process(es) must be engaged to facilitate the selection and execution of the correct task-relevant response.

This distinction between global and local control processes has been discussed in the literature on bilingual language control (e.g., de Groot & Christoffels, 2006) and fits well with models of bilingual language processing. In Green’s (1998) IC Model, language task schemas provide sustained control by orienting the system towards performing the goal-relevant task (i.e., speak in English), and potential responses that conflict with the current goals of the system are inhibited to prevent errors (i.e., Spanish lexicon). The top-down control of the language task schemas and the reactive inhibitory mechanism work together to resolve cross-language conflict between representations in the two languages. The revised Bilingual Interactive Activation (BIA+) model of word recognition (Dijkstra & van Heuven, 2002) also includes task schema and inhibitory mechanisms. Neurophysiologically-motivated models have postulated the importance of the frontal regions—including the anterior cingulate cortex and dorsolateral prefrontal cortex—to language control (Abutalebi & Green, 2008), implicating domain-general executive functions that support cognition more broadly (see also Levy & Anderson, 2002; Wang, Kuhl, Chen, & Dong, 2009).

Individual differences in global/sustained and local control likely contribute to multilingual speech production in different ways. In Green’s (1998) IC Model, sustained control may be supported by the maintenance of goal representations in WM—a critical component to the control of a range of goal-directed behaviors (e.g., Engle, 2002). Thus, individuals with larger WM capacity may better engage global control. In contrast, local control may be enacted by engaging inhibitory mechanisms, such as the reactive inhibition of representations that are competing with the target representation, as has also been posited by models of memory retrieval (e.g., Levy & Anderson, 2002). Green’s IC Model focused on reactive inhibition that is triggered by the activation of non-target representations. Colzato et al. (2008) identified a two-component model of cognitive control that included both a global control mechanism and a separate IC component. Colzato et al. also examined both reactive and proactive inhibitory mechanisms, which are likely to support bilingual language processes in different contexts.

IC and WM seem particularly relevant to language control during language switching. Indeed, a recent study found that better inhibitors showed smaller switch costs in a trilingual language switching task (Linck et al., 2012). These results parallel previous work demonstrating that language switching led to increased activation of the dorsolateral prefrontal cortex—specifically in areas previously linked to IC mechanisms responsible for resolving conflict (Hernandez, Dapretto, Mazziotta, & Bookheimer, 2001; Hernandez, Martinez, & Kohnert, 2000; for a review, see Abutalebi & Green, 2008).
Other recent work has demonstrated enhanced general executive functioning among bilinguals exhibiting better language control, on a range of cognitive control tasks requiring attentional control (Festman, Rodríguez-Fornells, & Münte, 2010) and conflict resolution (Festman & Münte, 2012).

These findings suggest that different executive control functions could be related to different components of language switching. Future work should be keen to probe this possibility. If IC is engaged to reduce cross-language representational conflict during speech production in a mixed language context, then IC should be most relevant to performance in conditions where the greatest amount of cross-language interference is expected: namely, when switching into or out of the L1. Specifically, it could be possible that better IC abilities should allow more efficient deployment of inhibition in the face of conflict that arises when switching into a previously irrelevant language, and thus should be related to smaller switch costs, due to faster latencies on switch trials (Linck et al., 2012). When it comes to WM resources, because the L1 is the dominant response, future work should investigate whether there are greater demands on WM resources to activate and maintain the task schema for L2 or L3 naming relative to L1 naming. Under these circumstances, WM should be most relevant when naming in the less dominant languages and more efficient activation and sustainment of the L2 naming and L3 naming task schemas should facilitate switching into the L2 and L3, leading to smaller switch costs (see also Linck, Michael, Golonka, Twist, & Schwieter, 2015).

5. The language switching paradigm

The language switching paradigm has been used extensively to argue for and against inhibitory control processes in bilingual speech production (Costa & Santesteban, 2004; Meuter & Allport, 1999; Schwieter & Sunderman, 2008). In this experimental procedure, participants name items (usually standardized black and white line drawings or Arabic numerals) in their L1 and L2 depending on a cue (normally the color of the background screen). Each item that is named is either a nonswitch trial, an item named in the same language as its previous trial or a switch trial, an item named in the opposite language as its previous trial. The items are normally presented individually on a computer screen in lists ranging from 5-14 pictures in length. Figure 1 shows an example of a list of 10 pictures with two switch trials (i.e., on trials 4 and 8).
Figure 1. Example of a list of pictures used in a language switching task. Note: E represents picture to be named in English (which would instead be shown to the participant on a blue background) and S for Spanish (which would instead be shown to the participant on a yellow background).

The critical measure in the language switching task has been the switch cost, the calculation of subtracting mean latencies on nonswitch trials from switch trials. Perhaps the first study to measure switch costs was Meuter and Allport (1999) in which “reasonably proficient” bilinguals participated in a battery of numeral switching experiments. The results suggested that: switch trials are slower than nonswitch trials; nonswitch trials have faster RTs in the L1; L2 switch trials have faster RTs than their L1 counterparts (which suggests that a switch to L1 is more difficult); and RTs increase with each successive switch. These findings were interpreted such that:

‘Negative priming’ arises from the active inhibition of one of two mutually competing languages (language A) which then persists involuntarily into the processing of the next task (language B). Thus, for language production in a weaker L2, active suppression of the competitor (L1) is needed. On a subsequent switch trial, this L2 language set generates powerful interference with the intended L1 response. For production in L1, in contrast, little suppression of any competitor language may be needed (Meuter & Allport, 1999, p. 35).

Since Meuter and Allport’s (1999) seminal study, several studies have reported asymmetrical switch costs that demonstrate that it takes more time to switch into the more-dominant language compared to the less-dominant language (Costa & Santesteban, 2004; Filippi, Karaminis, & Thomas, 2013; Jackson, Swainson, Cunnington, & Jackson, 2001; Philipp, Gade, & Koch, 2007; Santesteban & Costa, 2016; Schwieter, 2013; Schwieter & Ferreira, 2016; Schwieter & Sunderman, 2008, 2011; Tarłowski, Wodniecka, & Marzecova, 2013; Verhoef, Roelofs, & Chwilla, 2009). Interestingly, for bilinguals with similar dominance in both languages, symmetrical
switch costs have been observed (Costa & Santesteban, 2004; Costa, Santesteban, & Ivanova, 2006; Martin et al., 2013; Santesteban & Costa, 2016). However, a question that has recently been asked is whether or not the language switching paradigm and language switch costs are an accurate measure of the involvement of inhibitory control in speech production (Bobb & Wodniecka, 2013). In the next section, we will discuss the somewhat controversial interpretation of the asymmetrical switch cost.

5.1. Does the asymmetrical cost associated with language switching really imply inhibition?

On an empirical level, it is rather consistent that proficiency\(^1\) and the relative balanced between the two languages are related to the asymmetry reported in the language switching paradigm. However, there appears to be disagreement as to what this implies theoretically. It may indeed be premature to conclude that the absence of asymmetry means the absence of inhibition in language production (Bobb & Wodniecka, 2013). In fact, Bobb and Wodniecka argue that there are at least two reasons why the lack of asymmetry in the language switching paradigm should not be taken as solid support for the lack of inhibition. First, most studies have found asymmetrical switch costs in both L1 and L2 and “the presence of any switch costs, even if symmetric across languages, indicates some form of processing costs related to having had to shift between two language sets, some of which may (although not necessarily) be due to an inhibitory mechanism.” The researchers argue that these processing costs could either be due to other mechanisms than inhibition (Philip et al., 2007) or that both languages receive equal amounts of inhibition (see also Gollan & Ferreira, 2009).

Second, Bobb and Wodniecka (2013) argue that the lack of asymmetrical switch costs in language switching should not be taken as evidence for the lack of inhibition because there may be more than one locus or type of inhibitory control. For instance, some prior work has looked at active inhibition and local reactive inhibition (Colzato et al., 2008) while other research has made distinctions between global from local inhibition (de Groot & Christoffels, 2006; Christoffels, et al., 2007). If indeed there exist several types of inhibitions, then it is likely that a task such as language switching would require a number of these, not all of which can be fully assessed exclusively through the language switching paradigm.

Festman and Schwieter (2015) also note that, as with any task switch, if one of the tasks is more difficult than the other, it will take longer to switch into the easier task (see also Koch et al., 2010). Theoretically speaking, easy tasks require stronger inhibition than difficult tasks, which is why it is expected that when switching into a more dominant language, larger switch costs are observed compared to switching into a weaker language. In other words, this suggests that the more dominant language requires greater inhibition, which must be overcome when switching from a less dominant to

\(^1\) Schwieter and Sunderman (2008) explore the notion of proficiency through lexical robustness, which they define as an aspect of global proficiency that taps into the size and strength of the lexicon. Lexical robustness operationalizes the greater automaticity of word retrieval that comes with the familiarity with and frequency of its access (see also Costa et al., 2006; Schwieter & Sunderman, 2009, 2011).
more dominant, thereby leading to greater RTs. Festman and Mosca (2016) along with previous work by Guo, Liu, Chen, and Li (2013) demonstrate that the more time is given to prepare participants for the task, the better they are in their naming and switching performance. It is clear that various factors such as preparation time, task difficulty, and the relative level of proficiency play a role in determining the extent of language switch cost asymmetry, but what is still cloudy is whether this can be truly measured (exclusively) through language switch costs. As the field moves forward, it will be important that experimental paradigms converge to assess the contribution of inhibition to bilingual speech production.

5.2. Language switching vs. task switching: A look at language control and executive control

How does language control relate to domain-general executive control (EC) processes? This is a question recently asked by Calabria, Hernández, Branzi, and Costa (2012). Essentially, there are two lines of thought on this: bilingual language control processes are fully subsidiary to EC; or bilingual language control processes are only partially subsidiary to EC. There seems to be more support for the notion that language control is at least partially subsidiary to EC (Calabria et al., 2012; Linck et al. 2012). In fact, a number of neuroimaging studies have reported the same neural regions being engaged in language switching (Price, Green, & Von Studnitz, 1999; Hernandez et al., 2000; Hernandez et al., 2001; see Hervais-Adelman, Moser-Mercer, & Golestani, 2011 for a review) and non-linguistic switching tasks (Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; Crone, Wendelken, Donohue, & Bunge, 2006).

To measure domain general EC, a number of experimental tasks have been adopted. In one example of a task switching experiment, participants view the digits 1 through 9 (excluding 5) presented within either a gray or a white square in the center of the screen. When the color of the square is white, the participant indicates on the response box whether the presented digit was odd or even (“odd/even task”). When the square is gray, the participant indicates on the response box whether the digit was lower than 5 or higher than 5 (“low/high task”). As in the language switching paradigm, in task switching a switch cost latency score is computed as the difference between switch trial and non-switch trial latencies.

In another example of a task switching experiment, three shapes (square, circle, or triangle) and three colors (green, blue, and red) are selected. The shape condition is combined with the color condition to have nine total stimuli (e.g., green square, blue square, etc.). Participants are shown three shapes at a time on a computer screen: two at the top of the screen and one at the bottom. Their task is to match the triangle on the bottom with one of the two on the top according to one of two criteria: color or shape. This criteria is indicated by a cue (‘COLOR’ or ‘SHAPE’). Again, a switch cost latency score is computed as the difference between switch trial and non-switch trial latencies. The task switch cost should be larger for the easier response (sorting by color) compared to the more difficult response (sorting by shape) (Calabria et al. 2012).
In Calabria et al. (2012), the researchers compared language switching and task switching performance among highly-proficient bilinguals and highly-proficient bilinguals with a weak L3. The results showed that both bilinguals and bilinguals learning an L3 had symmetrical switch costs in the language switching task but not in the non-linguistic switching task. Furthermore, the researchers observed differential switching patterns within the experiments such that in the language switching task, the asymmetry of switch costs changed across the blocks whereas in the non-linguistic switching task, the asymmetry was continuous present throughout the task. Calabria et al. argue that these results suggest that bilingual language control is not completely subsidiary to the domain-general EC system. Another study which identifies important constraints on the conditions under which domain-general EC supports language switching is Linck et al. (2012), although this study provides some evidence of a direct link between inhibitory control abilities and language switching capabilities. The findings from both Calabria et al. and Linck et al. suggest that language control in bilinguals is not completely subsidiary to the domain-general EC system. In fact, a recent study by Shell, Linck, and Slevc (2015) did not find support for inhibitory control mechanisms underlying language switching. They instead argued that language switching and within-language lexical competition resolution do not share inhibitory processes.

6. Cognitive and neurological effects of bilingualism

In recent years, there have been a number of proposals that have started to explore the broader issues of how bilingualism as a daily experience shapes the neural networks that support language (e.g., Green & Abutalebi, 2013, 2016; Kroll & Bialystok, 2013; for reviews, see Costa, & Sebastián-Gallés, 2014; García-Pentón, Yuriem, Costello, Duñabeitia, & Carreiras, 2016; Li, Legault, & Litcofsky, 2014). Of particular interest has been the bilingual advantage in cognitive control reported in several studies (see Bialystok, Craik, Green, & Gollan, 2009). Below we take a look at this debate and discuss the influences of bilingualism on the brain.

6.1. The bilingual advantage debate

There is a growing body of research suggesting that a lifetime of being a bilingual entails certain cognitive benefits (Bialystok, 2009; Colzato, Bajo, van den Wildenberg, Paolieri, Nieuwenhuis, La Heij & Hommel, 2008). Indeed, bilinguals have been reported to outperform monolinguals in a variety of abilities involved in EC (i.e., controlling attention, inhibiting distraction, and shifting between mental sets). For instance, studies comparing younger and older monolinguals and bilinguals on the Simon task (Bialystok, Craik, Klein, & Viswanathan, 2004; Salvatierra & Rosselli, 2011), ambiguous figures (Bialystok & Shapero, 2005), modified antisaccade task (Bialystok, Craik, & Ryan, 2006), Stroop task (Bialystok, Craik, & Luk, 2008), and switching task (Gold, Kim, Johnson, Kryscio, & Smith, 2013) have shown significantly better performance for bilinguals. This bilingual advantage in EC is generally attributed to the continuous need to manage more than one activated language system. Recent research explains these advantages as a product of particular cognitive activities, experiences, and lifestyles which have long-term benefits that help to maintain cognitive functioning and to protect
against the aging effects that are associated with healthy aging, a term known as cognitive reserve (see Bialystok & Craik, 2015 for a review; but see also Paap, Johnson, & Sawi, 2015 for a critique against the existence of a bilingual advantage).

One reported advantage among bilinguals that may interact with their language switching abilities is their superior task switching performance compared to monolinguals (Prior & MacWhinney, 2010). In a comparison of task switching performance, Prior and MacWhinney found that bilinguals had reduced switch costs compared to monolinguals. However, bilinguals did not differ from monolinguals in the cost associated with mixed-task vs. single-task blocks. One interpretation of this finding is that bilinguals have an advantage when it comes to switching between task schemas and that this efficiency helps them to overcome, to some degree, language switch costs. Further support for this possibility comes from Linck et al. (2012) in which participants switched between their L1 and two significantly weaker languages. Better IC (i.e., small Simon effect) was associated with reduced language switch costs but only when switching into or out of the L1, when IC is most needed. These results suggest that speakers who are better at inhibiting may be able to more quickly engage IC mechanisms that are needed to constrain L1 activation before it becomes very strong, thus reducing the inhibition that is needed for the L1.

7. Conclusion

In this chapter we reviewed the main literature on bilingual lexical selection and competition, which has been mainly devoted to understanding whether the two languages of bilinguals get activated and compete for selection during speech production, as well as how do bilinguals select words in the intended language and avoid intrusions from their other language. We reviewed some of the most relevant evidence showing that the two languages get activated, focusing in the cognate facilitation effects showed by bilinguals in many different naming tasks due to the co-activation of their non-target language. Due to the wide agreement about the co-activation of both languages, research is now mostly focused on determining whether the two languages of bilinguals compete for selection, and whether language selection relies on inhibitory control mechanisms. Evidence suggesting that bilinguals rely on inhibitory mechanisms has been provided by means of language mixing and language switching effects elicited in language switching tasks, as well as by L2-immersion effects. However, as mentioned in our review, doubts have been casted about whether language switching paradigms really reflect the use of inhibitory mechanisms (Bobb and Wodniecka, 2013; see also Baus, Branzi, & Costa, 2016), and future research needs to establish whether language mixing and switching costs can be considered an index of inhibitory processes. The fact that the cognate status of words seems to modulate some of these effects also casts doubts about the use of inhibitory control during lexical access, or at least about the locus at which it might be applied. L2-immersion effects do not seem to affect L1 naming of cognate words (Baus et al., 2013), and cognate words have been argued to facilitate and/or trigger language switching (Broersma & De Bot, 2006). This evidence suggests that, if cognates effects arise at phonological level of
representation, as suggested by Costa and colleagues (Costa et al., 2000; 2005), either
the lexical representations of cognate translations are inhibited at late stages (Kroll et
al., 2006) or are not inhibited (Costa et al., 1999). Future research needs to explore
whether cognate words are differently represented and stored in the bilingual mind as
compared to non-cognates (as suggested by Strijkers, 2016), as well as whether the
cognate status of words affects, or even overcomes, the bilingual’s language control
mechanisms, as suggested by the “cognate triggering hypothesis” (Broersma & De Bot,
2006).

Finally, we also reviewed and discussed recent research
Explore the role of other type of executive control measures such as WM
Finally, the bilingualism effects and executive control

Thought questions

Although the cognate status of words is by definition based on a phonological
relationship, might cognate facilitation effects be originated at semantic, lexical and/or
phonological levels of representation?

Are cognate words represented in the brain in a different way than non-cognate words,
as suggested by Strijkers (2016)?

Might cognate words modulate the language control mechanism used by bilinguals, and
if so, would bilingual speakers of pairs of languages with large number of cognates
need to rely on stronger language control mechanisms than speakers of pairs of
languages with small number of cognates?

To what extent do bilingual language control mechanisms rely on general executive
function mechanism?

Does bilingualism benefit general executive control at all?

Related internet sites

Itziar Laka: How do bilinguals control their languages.
http://mappingignorance.org/2013/11/18/how-do-bilinguals-control-their-languages/

Francois Grosjean: When Bilinguals Speak.
https://www.psychologytoday.com/blog/life-bilingual/201305/when-bilinguals-speak

Idoia Ros: Bilingual advantages, with a pinch of salt.
http://mappingignorance.org/2016/06/13/bilingual-advantages-pinchof-salt/

Suggested references


References


Kambanaros, M., Michaelides, M., & Grohmann, K. (in press). Cross-linguistic transfer effects after phonologically based cognate therapy in a case of multilingual specific


Paap K., Johnson H., & Sawi O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex 69*, 265-278.


Authors’ bios

**Mikel Santesteban** is a Ramón y Cajal Research Fellow in the Bilingual Mind research group at the University of the Basque Country (UPV/EHU) in Spain. His research has appeared in journals including: *Applied Psycholinguistics; Bilingualism: Language and Cognition; Brain and Language; Cognitiva; Frontiers in Psychology; Journal of Experimental Psychology: Learning, Memory, and Cognition; Journal of Memory and Language; Language, Cognition, and Neuroscience; and Trends in Cognitive Sciences.*

**John W. Schwieter** is an Associate Professor of Spanish and Linguistics and Faculty of Arts Teaching Scholar at Wilfrid Laurier University in Canada and a Visiting Professor of Applied Linguistics in the Centre for Applied Research and Outreach in Language Education at the University of Greenwich in England. His research has appeared in journals including: *Bilingualism: Language and Cognition; The Canadian Journal of Applied Linguistics; Diaspora, Indigenous, and Minority Education: An International Journal; Language Learning; Linguistic Approaches to Bilingualism; The Mental Lexicon; and The Spanish Journal of Applied Linguistics.* His recent edited books include: *The handbook of translation and cognition* (2017, Wiley-Blackwell); *Cognitive