Agència de Gestió d'Ajuts Universitaris i de Recerca
AGAUR

UNIÓ EUROPEA
Fons Social Europeu

Memòria justificativa de recerca de les beques predoctorals per a la formació de personal investigador (FI)

La memòria justificativa consta de les dues parts que venen a continuació:
1.- Dades bàsiques i resums
2.- Memòria del treball (informe científic)

Tots els camps són obligatoris

1.- Dades bàsiques i resums

Títol del projecte ha de sintetitzar la temàtica científica del vostre document.
On the origin and scope of the second language speech production disadvatage

Dades de l'investigador (benficiari de l'ajut)

Nom	Cognoms
Elin	Runnqvist

Correu electrònic
elin_runnquist@yahoo.es

Dades del director del projecte

Nom	Cognoms
Albert	Costa Martinez

Correu electrònic
costalbert@gmail.com

Dades de la universitat / centre al que s'està vinculat
Universitat de Barcelona

Número d'expedient

Paraules clau: cal que esmenteu cinc conceptes que defineixin el contingut de la vostra memòria.
Language production; Bilingualism; Cross-language interactivity; Language control; Semantic interference

Data de presentació de la justificació
15/10 2012

Nom i cognoms, i signatura del beneficiari/ària

Vist i plau del/de la director/a del projecte

Generalitat de Catalunya
Departament d'Economia i Coneixement

Resum en la llengua del projecte (màxim 300 paraules)
Los hablantes bilingües tienen un acceso al léxico más lento y menos robusto que los monolingües, incluso cuando hablan en su lengua materna y dominante. Este fenómeno, comúnmente llamado “la desventaja bilingüe” también se observa en hablantes de una segunda lengua en comparación con hablantes de una primera lengua. Una causa que posiblemente contribuya a estas desventajas es el uso de control inhibitorio durante la producción del lenguaje: la inhibición de palabras coactivadas de la lengua actualmente no en uso puede prevenir intrusiones de dicha lengua, pero al mismo tiempo ralentizar la producción del lenguaje. El primer objetivo de los estudios descritos en este informe era testear esta hipótesis mediante diferentes predicciones generadas por teorías de control inhibitorio del lenguaje. Un segundo objetivo era investigar la extensión de la desventaja bilingüe dentro y fuera de la producción de palabras aisladas, así como avanzar en el conocimiento de las variables que la modulan. En lo atingente al primer objetivo, la evidencia obtenida es incompatible con un control inhibitorio global, desafiando la idea de mecanismos específicos en el hablante bilingüe utilizados para la selección léxica. Esto implica que una explicación común para el control de lenguaje y la desventaja bilingüe en el acceso al léxico es poco plausible. En cuanto al segundo objetivo, los resultados muestran que (a) la desventaja bilingüe no tiene un impacto al acceso a la memoria; (b) la desventaja bilingüe extiende a la producción del habla conectada; y (c) similitudes entre lenguas a diferentes niveles de representación así como la frecuencia de uso son factores que modulan la desventaja bilingüe.

Resum en anglès(màxim 300 paraules)
Bilinguals show a slower and less robust lexical retrieval than monolinguals, even when they speak in their first and dominant language. This phenomenon, commonly referred to as “the bilingual disadvantage”, also applies to second language speakers relative to first language speakers. One possible contributing cause of such bilingual disadvantages is the recruitment of inhibitory language control: inhibiting co-activated and interfering words from the language not in use while speaking might prevent cross language intrusions from occurring but slow down speech production. The first goal of the studies reported here was to test this hypothesis through different predictions put forward by inhibitory accounts of bilingual language control. A second goal was to investigate the scope of the bilingual speech production disadvantage both within and beyond single word retrieval and to refine our knowledge of the variables that modulate it. Regarding the first goal, the evidence gathered is incompatible with globally inhibitory language control and challenges the idea of bilingual-specific mechanisms used for lexical selection. This entails that a unified explanation of bilingual language control and the bilingual disadvantages in lexical retrieval is not likely. Regarding the second goal, the results show that (a) the bilingual disadvantage does not have an impact on memory retrieval; (b) the bilingual disadvantage extends to connected speech; and (c) cross-language similarities at different levels of representation and frequency of use are factors that modulate the bilingual disadvantage.

2.- Memòria del treball (informe científic sense limitació de paraules). Pot incloure altres fitxers de qualsevol mena, no més grans de 10 MB cadascun d'ells.

CHAPTER 1: IS BILINGUAL LANGUAGE CONTROL INHIBITORY?

(Runnqvist, E., Strijkers, K., Alario, F-X., & Costa, A. (2012). Cumulative semantic interference is blind to language: Implications for models of bilingual speech production. *Journal of Memory and Language*, 66 (4), 850-869.).

Abstract

Several studies have shown that concepts spread activation to words of both of a bilingual's languages. Therefore, a central issue that needs to be clarified is how a bilingual manages to restrict his speech production to a single language. One influential proposal is that when speaking in one language, the other language is inhibited. An alternative hypothesis is that bilinguals focus only on the language that is relevant for communication. Here these proposals were tested in a series of experiments in which Spanish-Catalan bilinguals named pictures. Cumulative semantic interference (CSI) was used as a window into lexical processing and cross-linguistic interactions. Results revealed that CSI is present between languages with the same magnitude as within-languages. This result cannot be accounted for by any of the above-mentioned models without substantial modifications. Instead, they are suggestive of bilingual processing dynamics qualitatively similar to those of monolinguals.

Keywords: Bilingualism; Language control; Speech production; Semantic competitor effects; Cumulative semantic interference

Introduction

An issue that has received much attention during the last years in the psycholinguistic research of bilingualism refers to how speakers control their two languages during speech production. Since it is well known that both languages of the bilingual speaker become active in the course of verbalization (e.g., Colomé, 2001; Hoshino & Kroll, 2008; Kroll, Bobb, & Wodniecka, 2006; Thierry & Wu, 2007), the research has been focused on how bilinguals are able to restrict speech to words of the desired language while avoiding interference from the unintended language.

There have been many different proposals regarding the functioning of such bilingual language control. Some of these involve the assumption that words from the unintended language are inhibited (e.g., Green, 1998); others that activity from the unintended language is ignored (e.g., Costa et al., 1999; Costa & Caramazza, 1999), and yet others that preverbal language membership encoding is sufficient to ensure lexical selection from the intended language (e.g., La Heij, 2005; Finkbeiner, Gollan, & Caramazza,, 2006). Despite extensive research, no consensus has yet been reached regarding the specific control mechanism, probably because most evidence is more consistent with one or another hypothesis than directly supporting of a particular view. Hence, it is important to bring new evidence that is able to discriminate between different assumptions.

In this article, we mainly focus on and test a combination of assumptions that has been embraced by what is perhaps the most influential model of bilingual language control: the Inhibitory Control Model (henceforth ICM; Green, 1998). The two assumptions of this model that are relevant for the present purposes are (a) lexical selection involves competition both within and across languages, and (b) lexical selection in the desired language is achieved by globally inhibiting any active lexical representations of the unintended language. As we will describe in more detail below, a model that combines these two assumptions makes a straightforward prediction regarding semantic contextual effects in picture naming tasks, namely that semantic competitor effects should be abolished or at least reduced in contexts of language alternation. Thus, the ICM offers a very specific and testable hypothesis about how lexical effects should interact with language control.

In the five experiments reported in this article this hypothesis was tested. The evidence gathered is not consistent with a model subscribing to these two assumptions (at least in the way they are explicitly put forward). In the General Discussion we consider how the ICM might be modified to account for the observed results and we also discuss alternative views of bilingual language control.

Semantic Interference Effects, Lexical Competition and Global Inhibition

Semantic contextual effects have a long-standing tradition in their use to inform models of language production (e.g., Abdel Rahman & Melinger, 2009; Alario, Segui & Ferrand, 2000; Bloem & La Heij, 2003; Bloem, van den Boogaard, & La Heij, 2004; Damian & Bowers, 2003; Damian, Vigliocco, & Levelt, 2001; Glaser & Dungelhoff, 1984; La Heij, Dirkx, & Kramer, 1990; La Heij, Starreveld, & Steehouwer, 1993; Roelofs, 1992; Rosinski, Golinkoff, & Kukish, 1975; Schriefers, Meyer & Levelt, 1990; Starreveld & La Heij, 1995, 1996; Wheeldon & Monsell, 1994). Crucial in the present context is the observation that the speed with which a given object can be named is affected by whether or not objects from the same semantic category have been named before. For example, naming "cat" leads to slower naming latencies when it is preceded by a semantically related word (e.g., *dog*, book, finger, *cat*) than when it is not (e.g., star, book, finger, cat). This phenomenon has been widely argued to reveal that lexical representations compete to be selected for production. More concretely, it is assumed that when a speaker wants to verbalize a given message, not only the intended words but also related words become active due to spreading activation (e.g., Caramazza, 1997; Dell, 1986; Levelt, Roelofs, & Meyer, 1999; Roelofs, 1992). Additionally, many researchers suppose that the level of activity of related words will influence the speed with which the intended word can be selected for further processing, a tenet referred to as "lexical competition" (e.g., Levelt et al., 1999; Roelofs, 1992). After producing a word (e.g., "dog") it is assumed that it has a slightly increased level of activation for some time. Therefore, when a related word has to be produced (e.g., "cat"), the previously produced semantic competitor (e.g., "dog") will interfere with lexical selection of the target[1]. The relevant question in the context of bilingual speech production is what happens if a semantic competitor has been named previously in another language than the target (i.e., "dog" and "gato" [cat]).

As discussed above the ICM model holds two main assumptions that are relevant in the present context: (a) lexical selection is competitive both within and across languages and (b) lexical selection in the desired language is achieved by globally inhibiting any active lexical representations of the unintended language.

The first assumption entails that, in principle, detrimental semantic competitor effects should occur across languages (e.g., "perro" [dog] can interfere with the production of "cat"). However, such cross-linguistic semantic competitor effects are counteracted by the second assumption, namely the global inhibition exerted over the non-target language. This interaction between semantically driven lexical competition and globally inhibitory language control is going to be crucial for the present study, and thus we explain it in some detail below.

As mentioned above, the ICM implements bilingual language control through a mechanism that suppresses activity stemming from the language that is not relevant for communication. The way the ICM formalises such language inhibition is via lexicon external task-schemas, responsible for controlling output goals (e.g., "translate from language A to language B" or "speak in language A"). The task schemas operate by regulating the level of activation of the lexical system through

[1] It should be noted that alternative ways of accounting for hampering effects of semantic relatedness in speech production that do not involve lexical competition have been proposed in the literature (e.g., Finkbeiner & Caramazza, 2006; Janssen, Schirm, Mahon & Caramazza, 2008; Mahon, Costa, Peterson, Vargas & Caramazza, 2007, Miozzo & Caramazza, 2003; Navarrete, Mahon & Caramazza, 2010; Oppenheim, Dell, & Schwartz, 2010). However, for now we will focus on the accounts that assume lexical competition given that this is a general tenet of the ICM; when evaluating any property of this model it seems appropriate to respect its theoretical basis.

connections with the language tags that indicate the language membership of lemmas. In that manner, the task schema will project an inhibitory signal to the lexicon suppressing all items containing a language tag of the unintended language. This mechanism prevents interference of the unintended language stemming both from co-activated words and from words that remain active due to previous production. More specifically, when a given task is maintained (e.g., speak language A), the inhibitory mechanism suppresses co-activated words with incorrect language tags (language B); when there is a change of language (e.g., switch to language B), the inhibitory mechanism suppresses those lemmas with incorrect language tags that have an increased activation level due to previous production (language A). As a consequence of this implementation, inhibition will have a global effect on the unintended language, inhibiting any active representation that bears an incorrect language tag, irrespective of the semantic or linguistic status of the words.

Thus, the combination of the two assumptions of the ICM leads to a very clear prediction regarding semantic competitor effects following a change of language: Items that have been named previously in language A will not have sufficient activation to cause interference when naming related items in language B on subsequent trials (i.e., "dog" will not interfere with naming of "gato" [cat]). Importantly, this prediction follows naturally from the combination of the two basic assumptions of the ICM without requiring any additional complex reasoning chains. In fact, the prediction is explicitly stated as follows by Green (1998, p.75): "(…) if there is a change of language then any lemmas in the previously active language will become inhibited. In certain circumstances, this should lead to the abolition of both cross-language and within-language competitor priming". In sum, the particular manner in which global inhibition as a language control mechanism is implemented in the ICM, provides the explicit prediction that producing a word in language A can only interfere with production of a related word in language B prior to the inhibition applied by the task schemas (that is, at the moment when we change the response language – because the inhibition still has to be performed – but not after; see Lee & Williams, 2001, for a similar reasoning).

Admittedly, this prediction is extreme in the sense that semantic competitor effects are expected to be completely abolished. One could think of a weaker version of the ICM that still allows for some semantic competitor effects in contexts of language alternation, assuming that inhibition of lexical items does not involve a complete reset in activity. Nevertheless, and at the very least, what is clear is that the combination of lexical competition and global inhibitory control would predict a reduction in the magnitude of semantic competitor effects (meaning a larger interference effect of "dog" when naming "cat" than when naming "gato").

Few studies have capitalized on the detailed position of the ICM concerning the interaction of language control with lexical competition effects. Instead, most studies have focused on testing one of these two assumptions in isolation. For instance, support for cross-language lexical competition has been obtained with bilingual versions of the picture-word interference paradigm where it has been found that presenting a written distracter in language A that is semantically related to a picture that has to be named in language B increases naming latencies compared to an unrelated distracter (e.g., Costa et al., 1999; Costa & Caramazza, 1999; Ehri & Ryan, 1980; Hermans, 2004; Hermans, Bongaerts, de Bot, & Schreuder, 1998; Mägiste, 1984, 1985). However, this finding remains silent about whether or not this interference is resolved with inhibition, let alone global inhibition. Similarly, evidence from the language switching paradigm has been taken to support inhibition of task-schemas. Since according to the ICM the amount of inhibition applied is proportional to activation, switch costs for L2 to L1 should be greater than vice versa during language switching, and response latencies should be slower when returning to a previously used response language in the n-2 back paradigm. This is exactly the pattern that has been observed in the literature (e.g., Meuter & Allport, 1999; Costa & Santesteban, 2004; Philipp & Koch, 2009). However, and besides some inconsistencies in the findings from this paradigm (e.g., symmetrical switch costs in a number of situations Costa & Santesteban, 2004; Costa, Santesteban, & Ivanova, 2006; Christoffels, Firk & Schiller, 2007; Gollan & Ferreira, 2009; Verhoef, Roelofs, & Chwilla, 2009), it is not clear whether the inhibition of task-schemas (e.g., speak L1, speak L2) identified in these paradigms actually affects the lexicon. That is, this evidence for inhibition remains silent about putative inhibitory connections between task-schemas and language tagged lemmas in the lexicon. For these reasons, we believe that testing the ICM's prediction of the absence of semantic competitor effects after switching response language can contribute with important insights, given that this prediction targets an interaction between both properties (i.e., the effects of global inhibition on a specific effect within the lexicon).

To our knowledge, only two studies have aimed at testing the hypothesis that a change of language should abolish semantic effects. Lee and Williams (2001) conducted an experiment in which participants produced a series of three words in response to a definition in L1 and then named two pictures either in L1 or in L2. The word elicited by the second definition could be semantically related to one of the picture names. It was observed that such semantic relatedness produced an interference effect in the following naming contexts ("*snow* - chair - flower - *rain*" or "*snow* - chair - flower - *pluie* [rain]"). In contrast, when a change of language occurred prior to the production of the target picture name (e.g., "*snow* - chair - arbre [tree] -*rain*" or "*snow* - chair - arbre [tree] - *pluie* [rain]"), the semantic interference effect disappeared. That is, having produced a word previously only interfered with naming of a related picture as long as there was no switch of language in between the critical items. This set of results seemed to fit perfectly with the ICM, given that "According to the IC model, switching to French suppressed currently active English lemmas and so eliminated, in advance, from the competition for selection in naming the probe picture in French, the English lemma associated with the lexical concept evoked by the definition" (Green, 1998, p.75). In other words, the data pattern uncovered by Lee and Williams (2001) is not only consistent with the notion of control through language inhibition, it actually provided direct evidence for global inhibition as described in the ICM.

Unfortunately, in an attempt to replicate and extend these results of Lee and Williams (2001), Li and MacWhinney (2011) observed that out of three groups tested only one performed similarly to the participants of Lee and Williams (2001). Furthermore, the results observed in the other two groups were rather complex and did not follow any obvious pattern. This unpredicted lack of systematic results reduces the appeal of the findings of Lee and Williams (2001) and questions the sensitivity of the paradigm, which at least calls for more empirical testing before embracing the notion of global inhibition. Nevertheless, we do believe that the basic rationale of the study of Lee and Williams (2001) could be crucial for enlarging our understanding of bilingual language control. Therefore, the aim of the current study was to further test the mechanism of global language inhibition of the ICM through the prediction regarding the abolition of competitor priming in contexts of language switching with an arguably simpler (i.e., it only involves picture naming) and more sensitive paradigm (e.g., Alario & Moscoso del Prado Martín, 2010): the cumulative semantic interference (henceforth CSI) paradigm (e.g., Brown, 1981; Howard, Nickels, Coltheart, & Cole-Virtue, 2006).

The Cumulative Semantic Interference paradigm

CSI refers to the finding that when naming pictures that belong to the same semantic category (e.g., cat, dog, pig etc.), response times increase linearly every time a new member is named (e.g., Brown, 1981; Costa, Strjikers, Martin, & Thierry, 2009; Howard et al., 2006; Oppenheim, Dell, & Schwartz, 2010; Navarrete, Mahon, & Caramazza, 2010). The exact paradigm we will rely on is based on that introduced by Howard et al. (2006). In that study, several members of different categories were intermixed with a varying lag between them (e.g., the members of a given semantic category were separated by 2, 4, 6 or 8 intervening items). The crucial manipulations within the stream of pictures were semantic category membership and ordinal position. The authors observed that, independently of lag, naming latencies increased by approximately 20 ms each time a new picture of a given semantic category was named. This effect was interpreted as a consequence of three properties inherent to the dynamics of speech production: (a) lexical competition (the activation level of "cat" can influence the speed with which "dog" can be selected); (b) priming (producing a word causes persisting strengthening in the mapping from semantics to lexical units or a persisting increment in the resting level of activation of lexical units); and (c) spreading activation (when we want to produce a word, nodes corresponding to semantically related words also become active). It should be noted that some recent accounts for the CSI-effect do not involve lexical competition (e.g., Oppenheim, Dell & Schwartz, 2010; Navarrete, Mahon, & Caramazza, 2010). However, due to the intrinsically competitive nature of the ICM, we believe that the most straightforward predictions of this model are made under the assumption of competition for selection. Therefore, we will defer discussion of a non-competitive mechanism behind the CSI effect and how such framework could account for this phenomenon in the bilingual case to the General Discussion.

As discussed above, according to the ICM, persisting priming between languages and even within language should be counteracted by the inhibition applied to the non-target language following a language switch. Thus, the ICM predicts that the CSI effects typically observed in a sequence such as "*cat* – tree – hand – *dog* – flower – star – *horse*" should be canceled out – both within and between languages – in a language alternating sequence such as "*cat* - árbol [tree] - mano [hand] - *perro* [dog] - flower - star – *horse*". In Experiments 1 to 3 we aimed at testing this prediction.

Experiment 1: assessing the presence of CSI in the two languages of a bilingual in language invariant naming contexts

Before measuring potential CSI across languages, it was necessary to establish whether such an effect could be obtained at all in L2 and whether it could be obtained with a similar magnitude to the effect in L1. Any conclusions drawn from effects obtained between languages hinges upon this assumption of comparability[2]. Because of this we ran a first experiment aiming at replicating the results of Howard et al. (2006) for L1 and extending them to L2.

Method

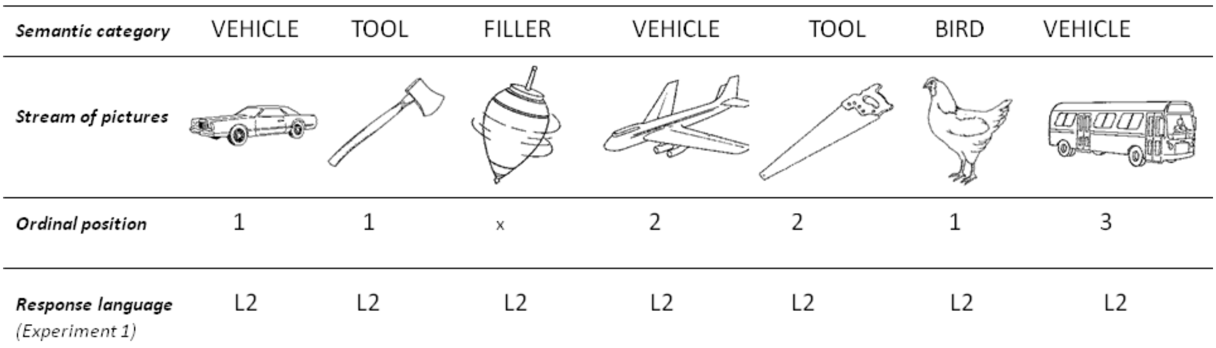
Participants. Two groups comprising a total of 48 of participants were tested. All participants were Spanish-Catalan bilinguals and undergraduate students of the University of Barcelona. Group 1 was comprised of 24 Spanish dominant (Spanish L1) speakers and group 2 of 24 Catalan dominant (Spanish L2) speakers (for more details regarding the participants of this and the following experiments, see Appendix A).

Design and procedure. The task of participants was to name a sequence of black and white line drawings (165 pictures repeated in two blocks). Of these pictures, 120 were target pictures and 45 were fillers (Appendix B contains a list of the target words used in this experiment). The 120 target pictures were five exemplars drawn from 24 different semantic categories embedded within the sequence and presented once in each block, and the crucial manipulation was the ordinal position in which each category member appeared. The lag between members of the same category was kept constant (2 intervening items) since Howard et al. (2006) showed that the number of intervening trials did not interact with the cumulative semantic interference effect. Hence, 11 items intervened between the first and last items in each category. The position in which a given category appeared within the sequence of pictures was counterbalanced across experimental lists by rotating the categories. Thus, 24 different lists were constructed so that each category appeared once in every possible position. 24 different combinations of two such lists were made in order for each subject to name two lists, and each of the resulting 24 combinations was named by one subject from each group. The reason we introduced a second block in the design was to increase power and reduce noise to optimize the between group comparison. No experimental item occurred in the first five pictures of each block. Items intervening between members of a category might either be drawn from other experimental categories or be fillers

[2] There are reasons why one might expect to find a reduced effect in L2. For example, words in L2 are presumably used less frequently than words in L1 making the links from concepts to words less robust (e.g., Kroll & Stewart, 1994; Gollan, Montoya, Cera, & Sandoval, 2008). This could arguably prevent activation from spreading to categorical members to the same extent as in L1. That is, while spreading activation among members such as “apple”, “orange” etc. of broad and/or frequent categories such as “fruit” might function similarly in L1 and L2; this might not be the case for members such as “lizard”, “chameleon” etc. of more fine-grained and/or less frequent categories such as “reptiles”.

(for an example of a sequence of trials see figure 1). Items within a set were also randomly ordered. Stimuli were black and white line drawings. The experiment was run in DMDX (Forster & Forster, 2003). Each picture was preceded by a visual cue for 500 ms, followed by a blank screen for 250 ms RTs were recorded by DMDX's voice key from the onset of the picture. The picture remained on the screen for 2000 ms, and was followed by a blank screen for 500 ms Subjects were instructed to name each picture with a single word as rapidly and accurately as possible.

Figure 1.



Analyses

Naming latency data. A common analysis protocol was followed for all experiments. Trials with errors were excluded (i.e., incorrect responses, disfluencies, voice key errors and omissions). The remaining naming latency data were fitted with linear regression models including crossed random and fixed effects (Baayen, Davidson, & Bates, 2008). A Box-Cox test (Box & Cox, 1964) indicated that a logarithmic transformation was the most appropriate to approximate a normal distribution (Experiments 1 through 5), and this is what was used. This transformation leads to a simple interpretation of the estimated beta values as percentage point deviation from baseline (intercept) performance. This is because expressing the model back in the natural response time scale requires using the exponential link function, and exp(beta) ~ 1 + beta for small beta. For further clarity, however, we report approximate estimates of effect sizes in milliseconds.

The fixed factors included in the models were always the same: trial number, ordinal position within each category, language spoken in the trial, as well as the interaction between ordinal position and language. Ordinal position was the key factor of theoretical interest. Language and the interaction were included to test whether the effects were equally true for L1 and L2. Finally, trial number was included to unconfound this variable from ordinal position in the category (Alario & Moscoso del Prado Martin, 2010).

The models included random factors (i.e., intercept estimates) for participants, items, and semantic categories. They also included mixed interactions between theoretically relevant fixed factors and the random factors for which they were treated within levels, thus ordinal position was crossed with participants, items and semantic categories (i.e., slope estimates). No other random factor was included. We followed Baayen's (2010) procedure of model criticism in which trials whose standardized residual value is above 2.5 are removed and the model recomputed. The tables always report the results of such recomputed models.

Error rates. The modelling procedure used to analyse error rates was similar to that used with response latencies, only now a binomial link function was used to predict the occurrence or absence of a correct response in every trial (Quene & van den Bergh, 2008), rather than its response time.

Results of Experiment 1

Naming latency data. The final model is summarized on Table 1. A total of 205 trials were excluded by the model criticism procedure. There were significant linear effects of trial number and ordinal position, and responses were significantly slower in L2 than in L1. There was no significant interaction between ordinal position and language.

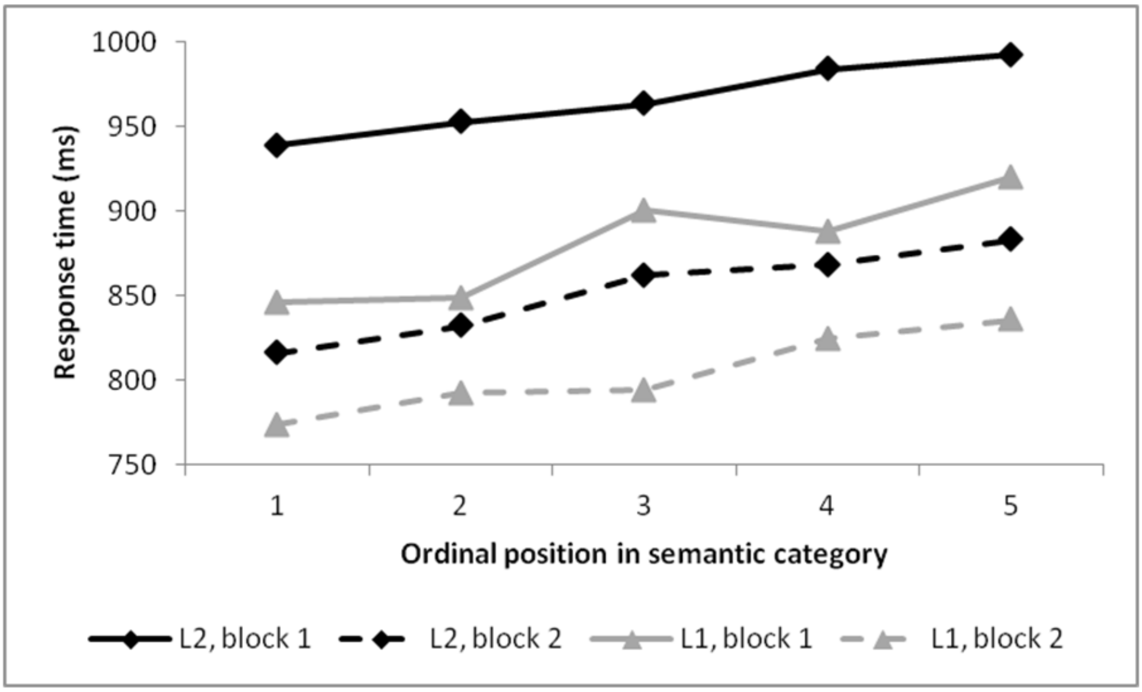
Table 1: Estimates of the linear regression model used to analyze the naming latency results of Experiment 1 (N trials = 9714)

	Beta	Approx. effect size (ms)	SE	T	P
Intercept	6.65	769.4	0.03	205.07	0

Trial number	-0.001	-0.4	0	-22.57	0
Ordinal position	0.02	18.4	0.003	7.62	0
Language (L2)	0.09	68.7	0.03	2.63	0.0043
Ordinal position*Language	-0.003	-2.6	0.003	-1.05	0.1462

Error rates. The average error rate was 14%. There was a significant effect of language on the probability of error; responses in L2 were more error prone than those in L1 [beta = .61, z = 2.10, p = .036]. There was no significant effect of ordinal position [beta = .06, z = 1.48, p = .139]. There was no significant interaction between ordinal position and language [beta = -.04, z = -.82, p = .41].

Figure 2.



Discussion

The results of Experiment 1 replicated those of Howard et al. (2006) and extended them to L2-production: regardless of dominance in the response language, naming a picture belonging to the same category as a previously named picture increased naming latencies by about 18 ms with each ordinal position (see Figure 2).

Importantly, the comparability of the magnitude of CSI in L1 and L2 validated the cross-language manipulation we wanted to introduce in the paradigm. Putative modulations of the effect in such a bilingual version could not be attributed to differences in L2 processing. However, in order to introduce two languages into the paradigm, several other modifications which we will specify below were necessary. Thus, Experiment 1 was not sufficient as a baseline because it could not rule out that any potential differences could be due to these other modifications. Therefore, in addition to the cross-language experiment (Experiment 2), we ran another experiment (Experiment 3) in which L1 and L2 were tested separately just as in Experiment 1, but using exactly the same design as that we wanted to employ for the cross-language experiment.

Experiments 2 and 3: CSI in language alternating versus language invariant naming contexts

The aim of these two experiments was to evaluate the presence of CSI effects in contexts where the response language is switched in between the critical items. In Experiment 2, the different instances of a given semantic category were named in different languages in an alternated fashion, while in Experiment 3 they were named in the same language. That is, we contrasted the effect of naming semantically related items in language-alternating conditions such as “cat, ….perro [dog], ….horse, ….vaca [cow]” against language-invariant conditions such as “cat, ….dog, ….horse, ….cow”.

The hypothesis regarding these two conditions according to the ICM is clear. If semantic contextual effects do not survive a switch of a response language, then we should not obtain a CSI effect in the language-alternating conditions. This is because language switching should lead to the inhibition of the lexical representations that are supposed to impair subsequent naming of items of the same semantic category. Accordingly, the language invariant condition (Experiment 3) should lead to a CSI effect but not the language alternating condition (Experiment 2).

There are some important modifications of the design in Experiment 2 compared to the previous experiment. First, we increased the number of items per semantic category from 5 to 10. In this way, five items were named in each language, allowing for the assessment of both cross-language and within-language CSI effects. Second, to reduce the potential effects of switching language on naming latencies, we decided to set up a design in which: a) the language switch was predictable, and b) the switch trial never corresponded to a target item. Hence, the critical target items (those that are used to calculate the CSI) were mixed with filler trials, resulting in a regular structure of targets and fillers throughout the whole experiment (i.e., filler-L1, filler-L1, target-L1, filler-L2, filler-L2, target-L2). In this way, target items were always placed in non-switch trials. Third, in order to be able to present fillers and targets in such regular sequence, we could only present one semantic category at the time (e.g., first all fruits, then all animals etc.) as opposed to intermixing members of different categories as in Howard et al. (2006). This is because in doing so, we already had a fixed lag of 2 intervening trials between languages and 6 intervening trials within languages. Although in the study by Howard et al. (2006) it was shown that lag did not interact with the effect, the maximum number of intervening trials in their experiment was 8, leaving the possibility open that a greater lag would attenuate the effect. There is certainly the possibility that all these modifications in the design beyond language switching might affect the magnitude of the CSI-effect. Hence, to be able to compare the magnitude in language alternating conditions against language-invariant conditions, we ran Experiment 3 that was identical in all respects to Experiment 2, but did not include language-switching. Rather, in Experiment 3 one response language was used throughout the whole experiment for a given participant (i.e., L1 or L2). To reiterate, we are interested in comparing the magnitude of CSI in a condition in which no language-switching is present (Experiment 3) vs. a condition in which the response language of the critical items is alternated (Experiment 2). Any difference between the two experiments will reveal the effects of language alternation on the semantic competitor effects. For the sake of clarity, we will present the results of the two experiments separately and subsequently we will compare the corresponding results.

Method

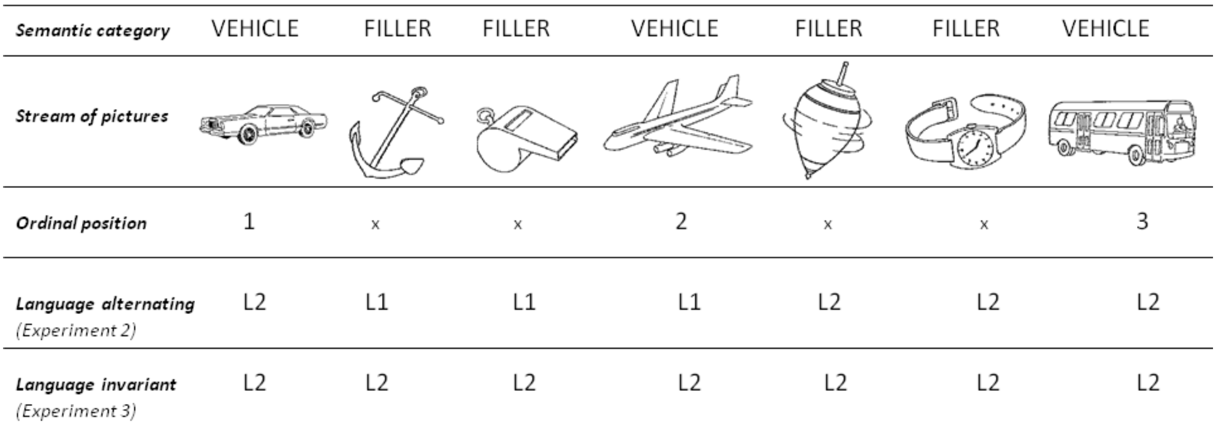
Participants. 60 undergraduate students of the University of Barcelona took part in Experiment 2. All participants were Spanish-Catalan bilinguals for whom Spanish was the first and dominant language. 30 participants began naming in Spanish (L1) and 30 in Catalan (L2).

56 undergraduate students of the University of Barcelona took part in Experiment 3. All participants were Spanish-Catalan bilinguals for whom Spanish was the first and dominant language. 30 participants completed the task in Spanish (L1) and 26 in Catalan (L2).

Design and procedure. The task of participants was to name a sequence of 309 black and white line drawings. Of these pictures, 100 were target pictures and 209 were fillers (Appendix B contains a list of the target words used for these and the following two experiments). The 100 target pictures were ten exemplars drawn from ten different semantic categories embedded within the sequence, and the crucial manipulation was the ordinal position in which each category member appeared. The lag between members of the same category was kept constant (2 intervening items) so that 26 items intervened between the first and last items in each category. Categories were presented in a clustered fashion (i.e., all ten members of a given category were presented before the first member of a new category could appear). Because of this, all items intervening between members of a category were fillers (for an example of a sequence of trials, see figure 3). No experimental item occurred in the first five pictures of each block. All pictures were presented with a coloured frame (red or blue) that varied every three pictures (on the trial immediately after a target trial, e.g. red filler, red filler, red target, blue filler etc.) so that in each category five members were presented with a blue frame and five with a red frame. Subjects in Experiment 2 were instructed to choose naming language according to the colour of the picture frame (blue or red). The correspondence between colour and language was counterbalanced across participants so that for 50% red indicated Spanish and blue Catalan while for the other 50% blue indicated Spanish and red Catalan. In experiment 3 the same coloured picture frames were used, but participants were instructed to ignore them (since there was only one response language). The frame appeared 500 ms prior to stimulus onset. 30 experimental lists were created: in Experiment 2 half of the participants started naming in L1 and the other half in L2; in Experiment 3 half of the participants named the whole list in L1 and half of the participants in L2. The lists were constructed by creating three initial lists in which members were randomly assigned a position within their category and categories were randomly assigned a position within the experiment with the already mentioned restrictions of lag and clustered appearance of categories. In each of these three lists, position of items within a category (1-10) was rotated so that all category members appeared equally often in each ordinal position, resulting in 30 different lists. Across these lists, the position in which a given category appeared within the sequence of pictures was also counterbalanced by rotating the categories. The experiment was run in DMDX (Forster & Forster, 2003). Each picture was preceded by a visual cue for 500 ms, followed by a blank screen for 250 ms RTs were recorded by DMDX's voice key from the onset of the picture. The picture remained on the screen for 2000 ms, and was followed by a blank screen for 500 ms

Subjects were instructed to name each picture with a single word as rapidly and accurately as possible.

Figure 3:



Results of Experiment 2: language alternating naming context

Naming latency data. The final model is summarized on Table 2. A total of 61 trials were excluded by the model criticism procedure. There were significant linear effects of trial number and ordinal position, and responses were significantly slower in L2 than in L1. There was no significant interaction between ordinal position and language.

Table 2: Estimates of the linear regression model used to analyze the naming latency results of Experiment 2 (N trials = 4486)

	Beta	Approx. effect size (ms)	SE	T	P
Intercept	6.924	1016.4	0.04	195.89	0

Trial number	0	0.1	0	3.77	1e-04
Ordinal position	0.012	12.1	0.002	5.74	0
Language (L2)	0.034	34.9	0.01	2.46	0.007
Ordinal position*Language	-0.001	-1.5	0.002	-0.66	0.2563

Error rates. Error-rates are summarized on Table 6. There was no significant effect of language on the probability of error [beta = .21, z = 1.36, p = .175]. There was a significant effect of ordinal position; error probability increased with the position within the category [beta = .05, z = 2.39, p = .017]. There was no significant interaction between ordinal position and language [beta = .00, z = .06, p = .949].

The results of Experiment 2 thus indicate that the CSI effect is present across the 10 ordinal positions even when the different exemplars of a given semantic category are named in a language alternating fashion. Put differently, language switching does not remove the CSI effect. However, before understanding whether the language switching actually has any effect at all on the magnitude of CSI we need to assess its magnitude in language-invariant conditions.

Results of Experiment 3: language invariant naming context

Naming latency data. The data were analyzed as in the previous experiment and the final model is summarized on Table 3. A total of 78 trials were excluded in the model criticism procedure. There were significant linear effects of trial number and ordinal position, and responses were significantly slower in L2 than in L1. There was no significant interaction between ordinal position and language.

Table 3: Estimates of the linear regression model used to analyze the naming latency results of Experiment 3 (N trials = 4280)

	Beta	Approx. effect size (ms)	SE	T	P
Intercept	6.807	904.6	0.04	165.05	0
Trial number	0	0.1	0	3.31	5e-04

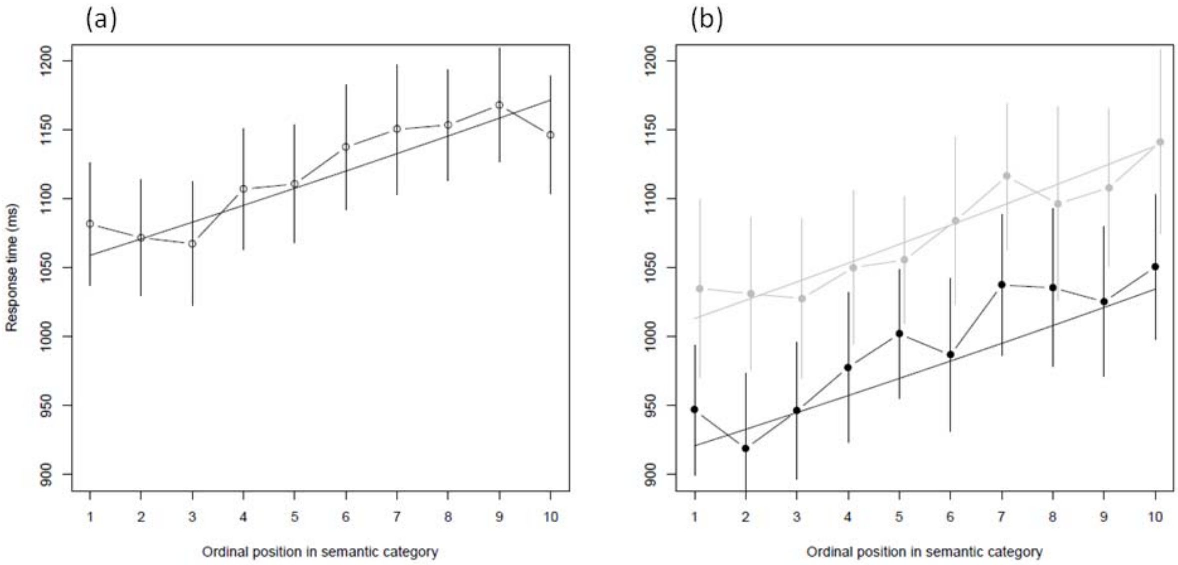
Ordinal position	0.014	12.5	0.002	6.06	0
Language (L2)	0.107	96.9	0.04	3.11	0.001
Ordinal position*Language	-0.002	-1.9	0.002	-0.87	0.1922

Error rates. Error-rates are summarized on Table 6. There was a significant effect of language on the probability of error; responses in L2 were more error prone than those in L1 [beta = .61, z = 2.64, p = .008]. There was a significant effect of ordinal position; error probability increased with the position within the category [beta = .05, z = 2.27, p = .023]. The interaction between ordinal position and language was not significant [beta = .00, z = .01, p = .993].

Combined Analysis of Experiments 2 and 3

A crucial test of the effects of language alternation on the presence of the CSI was conducted by testing the significance of the interaction between the factor Experiment (Experiment 2 vs. Experiment 3) and the linear estimates computed over the combined dataset. There was a main effect of Experiment, whereby participants were faster in Experiment 3 than in Experiment 2 [beta = .087, t(8750) = 3.41, p = .000]. In other words, at least in this between-participant comparison, language alternation incurs into a global RT cost. Not surprisingly, there was also a significant effect of ordinal position [beta = .013, t(8750) = 6.61, p = .000]. Crucially, the interaction between ordinal position and Experiment was not significant [t(8750) = -0.80, p = .212]. This reveals that the CSI effect is statistically undistinguishable across experiments (see figure 4). In keeping with the absence of interaction just noted, the approximate effect sizes for ordinal position in the category, reported in Tables 2 and 3, were very similar across experiments.

Figure 4.



Discussion

The main objective of these two experiments was to assess how language alternation affects a contextual semantic effect in object naming, namely the CSI effect. To do so, we evaluated the CSI effect in two conditions. In Experiment 2, the exemplars of a given semantic category were named in two different languages in an alternating fashion. In contrast, Experiment 3 only involved one language. The results of the two experiments were clear (see figure 4). First, the CSI effect was present in language invariant conditions (Experiment 3) as well as in language alternating conditions (Experiment 2). Importantly, the magnitude of the CSI effect was undistinguishable in both experiments, indexing that it is not sensitive to naming context (i.e., language alternating or invariant).

These results have important implications for one of the basic tenets of the ICM. The ICM assumes that producing words in a given language involves the global inhibition of any active lexical items of the other language. When there is a language switch the active lexical representations of the just abandoned language are inhibited, presumably rendering their activation level low. As Green (1998) argues, in this way, semantic contextual effects should not survive a language switch. Hence, the presence of a CSI effect in Experiment 2 is clearly at odds with the assumption of global inhibition, at least as implemented in the ICM.

Nonetheless, it should be acknowledged that the ICM is ambiguous as to what extent persisting activation from the non-target language is inhibited, leaving open the possibility that the suppression is not complete. In this way, one could argue that the prediction regarding the abolishment of semantic competitor effects is too extreme, since such abolishment is not a necessary consequence of this model. If inhibition of the residual activation of the non-target language is not complete, then it is still possible to observe semantic competitor effects across languages while preserving the mechanism of global inhibition. However, even when adopting such weaker version of the ICM, these cross-language effects should at the very least be of a smaller magnitude than semantic competitor effects in a language invariant context. This is because even if the inhibition does not involve a complete suppression of activity, by definition it does involve a lowering of activity. Hence, the activation that survives this inhibition should be lower than if no inhibition would have been applied (as in the language invariant context), leading to a reduced magnitude of interference effects. This is not what we observed in our experiments, where the magnitude of CSI was equal in language alternating and language invariant contexts.

Importantly, these data are also problematic for other models of bilingual language control. Although no other proposal emulates the ICM in terms of specificity and consequently testable predictions, another recurrent way of solving the issue of bilingual language control has been by eliminating cross-linguistic lexical competition and hence eliminating the need of inhibition (i.e. the language specific selection hypothesis by Costa and colleagues, e.g., Costa et al., 1999, Costa & Caramazza, 1999). In such a view, after parallel activation of the two languages, the lexical selection mechanism is thought to be language-specific. That is, a lexicon external device is responsible for programming which of the two lexicons is to be considered for selection, and only lexical nodes pertaining to that lexicon will compete as possible output candidates. According to this view, CSI effects in language alternating conditions should not be present with the same magnitude as in language invariant conditions. In fact, the language-specific selection hypothesis will predict the presence of two parallel cumulative effects in the language alternating conditions, one for each language. This is because the lexical activation of language A will never compete with that of language B. Consequently, semantic interference should not accumulate across languages, but only within languages. That is, in a condition such as “cat, ….perro [dog], ….horse, ….vaca [cow]” there will be two parallel CSI-effects, one including “cat” and ”horse” and the other including “perro” and ”vaca”. Accordingly, this parallel CSI effect for each language would predict a difference in the slope of the CSI effect when compared against the language invariant condition “cat, ….dog, ….horse, ….cow”. Strictly speaking, the magnitude of the CSI effect in the language alternating condition should be half of that of the language invariant condition. This prediction is clearly at odds with the results of Experiments 2 and 3.

Note, however, that a different argument can be made in favour of the ICM and the language specific selection hypothesis, because we do not know precisely how language mixing may interact with the CSI effect. Perhaps because of the increased difficulty of the task induced by the language mixing, any lexico-semantic competition effect increases in magnitude. If this were to be the case, differences in the magnitude of the CSI effect in the language invariant and language alternating conditions might be masked by a language mixing effect. The results obtained in Experiment 2 would then not be conclusive against a weak version of the ICM or the language-specific selection hypothesis.

In short, the results of Experiment 2 may reflect attenuated CSI effects occurring either across languages or within each language separately, but re-boosted by the mixing of languages. If this were the case, then a string such as “cat, ….manzana [apple], ….horse, ….pera [pear]” in which two undoubtedly independent semantic accumulations occur in each language (one for animals, one for fruits) should elicit similar results to a string such as “cat, ….perro [dog], ….horse, ….vaca [cow]”, since in both examples the amount of (boosting) language mix is equal. Alternatively, a reduced CSI effect in the former example compared to the latter would signal that the accumulation occurs across the two sets of animals, and hence across languages. Experiments 4 and 5 were designed to contrast these possibilities.

Experiments 4 and 5: on the origin of the CSI in language alternating naming contexts

Experiment 4 was identical to Experiment 2, except that half of the items corresponding to language A were replaced by items of a different semantic category (e.g., members of fruits in language A alternating with members of animals in language B). Thus, just like Experiment 2, Experiment 4 involved language switching, but critically only half of the ordinal positions 1-10 (odd vs. even) were actually members of the same semantic category. This design allows testing the alternative described above.

However, because we could not know if the nested design would modulate the effect, Experiment 3 could no longer be considered as a reliable baseline against which to compare the effects of Experiment 4. Therefore, the goal of Experiment 5 was to obtain a baseline for the nested appearance of semantic categories: the design was identical to that of Experiment 4 except that only one response language was used throughout the experiment (50% of the participants named in Spanish and 50% in Catalan).

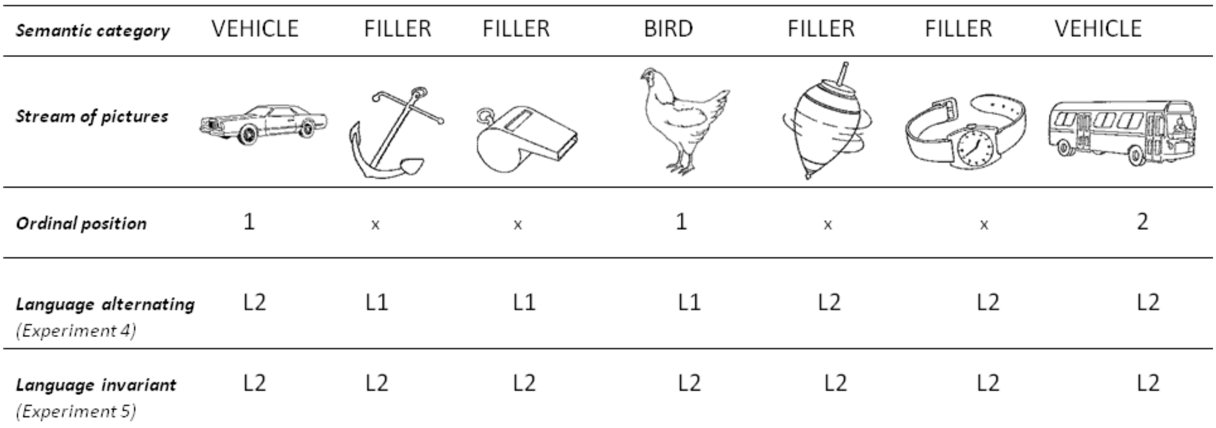
In short, the critical comparison to better understand the pattern of CSI across languages is that between the magnitude of the CSI effect in Experiment 2 and in the present Experiment 4. Any difference between the two experiments would indicate the contribution of authentic cross-language effects to the magnitude of the CSI in Experiment 2. Before that, however, the comparison between Experiments 4 and 5 will clarify whether or not changes in the design (i.e., the nested appearance of semantic categories) influence the magnitude of the CSI effect.

Method

Participants. 60 undergraduate students of the University of Barcelona took part in each of the experiments. All participants were Spanish-Catalan bilinguals for whom Spanish was the first and dominant language. In Experiment 4, 30 participants began naming in Spanish (L1) and 30 in Catalan (L2). In Experiment 5, 30 participants named the pictures in Spanish (L1) and the other 30 in Catalan (L2).

Design and procedure. Unless stated otherwise, everything in the design and procedure was identical to experiments 2 and 3 (for an example of a sequence of trials see figure 5). 30 experimental lists were created by modifying the lists of Experiment 1: the five members corresponding to ordinal positions 1, 3, 5, 7 and 9 in experiments 1 and 2 were moved to replace the members of the same ordinal positions two categories further. In Experiment 4, subjects were instructed to choose naming language according to the colour of the picture frame (blue or red). The correspondence between colour and language was counterbalanced across participants so that for 50% red indicated Spanish and blue Catalan while for the other 50% blue indicated Spanish and red Catalan. In Experiment 5 participants named all the pictures in one language (either Spanish or Catalan).

Figure 5.



Results of Experiment 4: language alternating naming context

Naming latency data. The data were analyzed as those of the previous experiments, and the final model is summarized on Table 4. A total of 70 trials were excluded in the model criticism procedure. There was a significant linear effect of trial number and ordinal position. Responses were slower in L2 than in L1, but there was no significant interaction between ordinal position and language.

Table 4: Estimates of the linear regression model(s) used to analyze the naming latency results of Experiment 4 (N trials = 4253)

	Beta	Approx. effect size (ms)	SE	T	P
Intercept	6.928	1020.9	0.04	175.90	0

Trial number	0	0.2	0	6.41	0
Ordinal position	0.007	6.9	0.002	3.13	9e-04
Language (L2)	0.073	74.6	0.01	5.12	0
Ordinal position*Language	0	-0.1	0.002	-0.03	0.4884

Error rates. Error-rates are summarized in Table 6. There was a significant effect of language on the probability of error; responses in L2 were more error prone than those in L1 [beta = .64, z = 4.06, p = 0]. There was no significant effect of ordinal position [beta = .02, z =.97, p = .332], nor an interaction between ordinal position and language [beta = -.01, z = -.23, p = .819].

Results of Experiment 5: language invariant naming context

Naming latency data. The data were analyzed as those of the previous experiment, and the final model is summarized on Table 5. A total of 70 trials were excluded in the model criticism procedure. There were significant effects of trial number and ordinal position. For the first time, there was not a significant effect of language. Finally, there was no interaction between ordinal position and language.

Table 5: Estimates of the linear regression model(s) used to analyze the naming latency results of Experiment 5 (N trials = 4195)

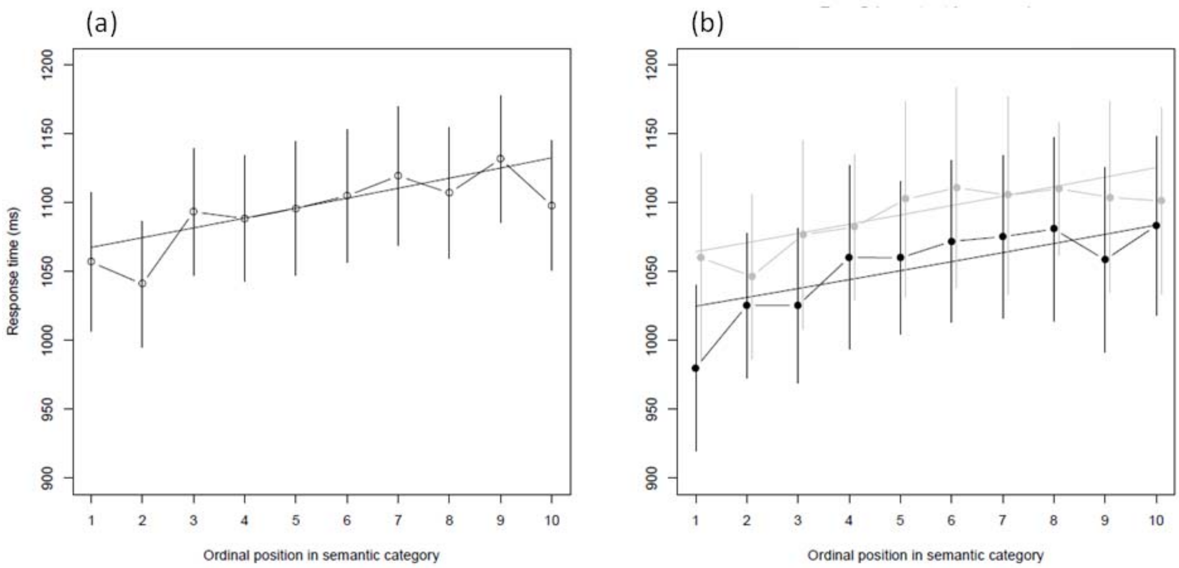
	Beta	Approx. effect size (ms)	SE	T	P
Intercept	6.930	1022.6	0.05	155.18	0
Trial number	0	0.2	0	4.88	0
Ordinal position	0.007	7.5	0.002	4.28	0
Language (L2)	0.050	55	0.04	1.36	0.0873
Ordinal position*Language	-0.003	-3.1	0.002	-1.38	0.0845

Error rates. Error-rates are summarized on Table 5. There were no significant effects on the error rates.

Combined analysis of Experiments 4 and 5

We then combined the data from Experiments 4 and 5, and tested explicitly whether the CSI-effect differed across the two experiments. This was done by testing the significance of the interaction between the factor Experiment and the cumulative effect computed over the combined dataset. There was no main effect of Experiment [beta = -0.018, t(8446) = -0.61, p = .2708]. The pattern for the other main predictors revealed a significant effect of ordinal position [beta = 0.006, t(8446) = 4.22, p = .000] and a significant effect of language [beta = 0.073, t(8446) = 11.85, p = .000] . The interaction between ordinal position and Experiment was not significant [t(8446) = -0.24, p = .406]. This pattern is summarized on Fig. 6.

Figure 6.



Combined analysis of Experiments 2 and 4

Naming latency data. We compared Experiments 2 and 4 within a single analysis, to estimate whether the CSI effect differed between these experiments. There was no significant effect of

Agència de Gestió d'Ajuts Universitaris i de Recerca AGAUR

UNIÓ EUROPEA Fons Social Europeu

experiment [beta = 0.0179, t(8734) = .6467, p = .2589]. There was a main effect of ordinal position [beta = -0.0115, t(8734) = 6.8520, p = .000]. More importantly, there was a significant interaction between ordinal position and experiment [beta = -0.0049, t(8734) = -3.0147, p = .0013], indicating significantly smaller CSI in Experiment 4 than in Experiment 2.

Discussion

The aim of Experiments 4 and 5 was to shed some light on the contribution of authentic cross-language effects to the CSI observed in Experiment 2. In particular, we wanted to assess whether the magnitude of the CSI effect observed in Experiment 2 was due to the difficulty added to that experiment by including language mixing. To do so, in Experiment 4 half of the items corresponding to language A were replaced by items of a different semantic category (e.g., members of fruits in language A alternating with members of animals in language B).Thus, just like Experiment 2, Experiment 4 involved language mixing, but critically only half of the ordinal positions 1-10 (e.g., 1, 3, 5, 7 and 9) were actually members of the same semantic category. To test the impact of this nested design in isolation and obtain a baseline for Experiment 4, we ran Experiment 5 in which the same design was used but responses came only from one language (i.e., either Spanish or Catalan).

The results showed that the magnitude of the CSI effect in Experiment 4 was identical to that of Experiment 5 and reduced in comparison to Experiment 2, meaning that whether or not the experiment included language alternation was irrelevant for the CSI effect size. That is, the CSI effect in conditions such as “cat, ….manzana [apple], ….horse, ….pera [pear]” was approximately half the size compared to conditions such as “cat, ….perro [dog], ….horse, ….vaca [cow]”. As only the latter condition included semantic relatedness across languages, this interaction critically reveals the contribution of authentic cross-language effects to the magnitude of CSI.

Before moving on to the General Discussion, we would like to dedicate some words to a possible caveat of our study related to the similarity of the languages that we tested. Given the high overlap in form between Spanish and Catalan (which is reflected by the proportion of cognates – approximately 2/3– used in the current study), one might wonder whether these results would be replicated with a different combination of languages with less formal overlap between translation words (a bigger non-cognate proportion). In other words, are these results, and in particular the identical CSI magnitude in language alternating and language invariant naming contexts, induced by

Generalitat de Catalunya
Departament d’Economia
i Coneixement

cognates? To answer this question, we conducted additional analyses exploring the potential contribution of cognate status to the CSI effect across languages. We restricted these comparisons to Experiments 2 and 3 since these two experiments were equal in everything except for the variable of interest: between language versus within language CSI. Specifically, we estimated CSI effect per category, namely the mixed interaction between the fixed effect 'ordinal position' and the random effect 'semantic category'. We then examined whether the effect varied systematically as a function of the amount of cognates present across categories. In the absence of a definite definition, two different measures of cognate distribution were tested: (a) the percentage of cognates within each category; and (b) the contrast between the five categories with fewer cognates (vegetables, furniture, tools, birds and clothes) versus the five categories with most cognates (musical instruments, vehicles, electrical appliances, audiovisual equipment and zoo animals) among a total of ten categories. Analysis of variance (Anova) were used because the dependent variable (random CSI effect per category) is normally distributed by construction. A contribution of cognate status to the CSI effect across languages would surface as an interaction between experiment and either measure of cognate status.

This interaction did not approach significance in neither of the two analyses (analysis 1: $F(1, 8) = 2.788$, $MSE = 1.695$, $p = .134$; analysis 2: $F<1$). Moreover, as can be appreciated in figure 7, the pattern in the magnitude of CSI was rather similar across categories between Experiment 2 and Experiment 3. In any event, we acknowledge that there is still a small possibility that the results reported here may not apply to bilinguals who speak a combination of less similar languages.

Figure 7.

Agència de Gestió d'Ajuts Universitaris i de Recerca
AGAUR

UNIÓ EUROPEA
Fons Social Europeu

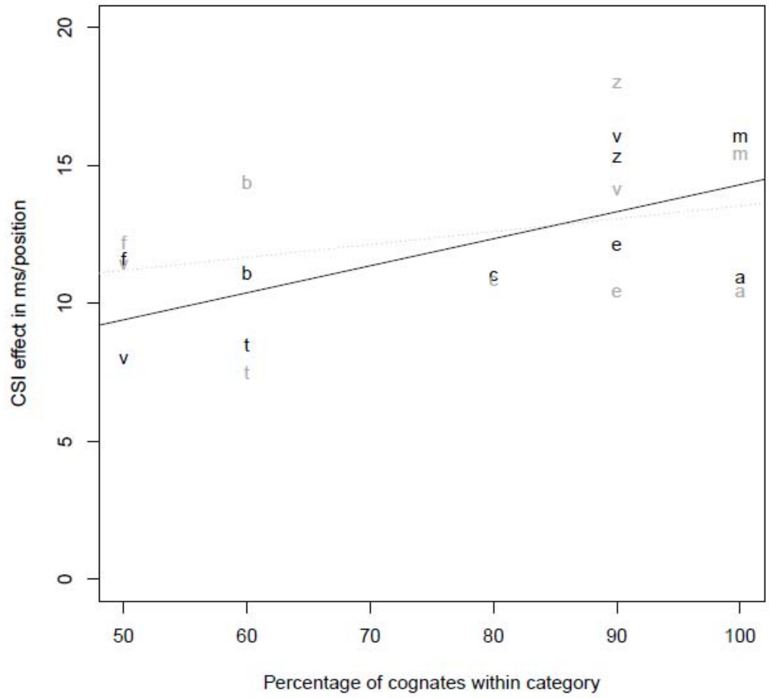


Table 6: Error-rates in percentages of experiments 2-5 broken down by ordinal position within semantic category and response language.

Exp 2			Exp 3			Exp 4			Exp 5		
ord.pos	lang	Err	ord.pos	lang	Err	ord.pos	Lan g	Err	ord.pos	lang	Err
1	L1	18.7	1	L1	18.7	1	L1	26.0	1	L1	30.3
2	L1	20.3	2	L1	17.3	2	L1	21.7	2	L1	26.0
3	L1	20.7	3	L1	20.7	3	L1	26.0	3	L1	28.0
4	L1	22.3	4	L1	16.3	4	L1	19.3	4	L1	28.3
5	L1	21.0	5	L1	16.7	5	L1	26.0	5	L1	23.7
6	L1	26.7	6	L1	20.0	6	L1	21.0	6	L1	27.0
7	L1	24.0	7	L1	16.0	7	L1	23.7	7	L1	28.0
8	L1	23.0	8	L1	18.3	8	L1	27.7	8	L1	27.3
9	L1	24.7	9	L1	20.0	9	L1	27.7	9	L1	30.3
10	L1	25.7	10	L1	22.7	10	L1	21.0	10	L1	26.3
1	L2	22.7	1	L2	25.0	1	L2	27.7	1	L2	29.3
2	L2	22.7	2	L2	25.8	2	L2	37.7	2	L2	30.0
3	L2	24.0	3	L2	24.6	3	L2	28.0	3	L2	31.0
4	L2	25.0	4	L2	24.6	4	L2	32.0	4	L2	33.0
5	L2	27.3	5	L2	26.5	5	L2	26.0	5	L2	29.3
6	L2	24.3	6	L2	25.0	6	L2	36.3	6	L2	28.7
7	L2	26.0	7	L2	25.8	7	L2	32.0	7	L2	28.0
8	L2	29.0	8	L2	26.9	8	L2	34.3	8	L2	31.0
9	L2	26.7	9	L2	27.3	9	L2	31.7	9	L2	29.7
10	L2	29.7	10	L2	30.8	10	L2	33.3	10	L2	33.0

Generalitat de Catalunya
Departament d'Economia
i Coneixement

General Discussion

The goal of the present study was to advance in our knowledge about how bilinguals manage to restrict their speech to only one language. More concretely, we tested the predictions made by a model that embraces the following two assumptions: (1) Lexical selection involves competition within and across languages, and (2) it involves the inhibition of any active lexical items of the unintended language. Following the proponent of this model (Green, 1998), we argued that the combination of these assumptions leads to the prediction that semantic competitor effects will be absent or reduced in a task that involves language alternation compared to a task that does not. We aimed at testing this prediction through five experiments involving more than 250 Spanish-Catalan bilingual participants, employing a bilingual version of the CSI paradigm in which the response language of the semantically related words was either alternated or blocked. This design also allowed us to test the predictions of another proposal of bilingual language control according to which (a) words only compete within languages; and (b) no inhibition is applied to words belonging to the unintended language (e.g., Costa et al., 1999). We argued that these assumptions lead to the prediction that the CSI effect should be halved when language is alternated among the critical items compared to when the same language is used for all items. Thus, we explored the presence and magnitude of semantic competitor effects in various conditions, paying special attention to whether or not it was affected by language-alternation. The following main results were observed:

1) CSI effects are similar in L1 and L2 (Experiment 1).
2) CSI effects are of similar magnitude across (Experiment 2) and within (Experiment 3) languages, which means that alternating response-language does not reduce the effect.
3) The magnitude of the CSI effect is not affected by language mixing *per se* (Experiments 2, 4 and 5).
4) CSI is stronger across members of the same semantic category named in two different languages than across members of different categories (where it is absent; Experiment 2 vs. Experiment 4); hence there is genuine cross-language accumulation

The combination of these four observations allows for the following empirical generalization: Semantic competitor effects are insensitive to language membership (L1 or L2) and to language alternation. Currently, there is no model of bilingual language control specified in the literature that can

Agència de Gestió d'Ajuts Universitaris i de Recerca
AGAUR
UNIÓ EUROPEA
Fons Social Europeu

account for these findings. In the remainder of this paper we will discuss this claim in more detail and point out alternative ways of integrating our results with the existing literature. We will begin by briefly refreshing the basic assumptions of Howard et al.'s account of CSI and also introduce those of an alternative account of CSI put forward by Oppenheim and colleagues (2010). These accounts in combination with the findings from the current study will then be used as constraints when discussing why existing proposals of bilingual language control are unable to accommodate our findings and how they should be modified in order to do so. Finally, we will argue that the most parsimonious account for the results of the current study consists in assuming that monolingual and bilingual speech production are qualitatively similar since mechanisms that already exist within the monolingual speech production architecture may be enough to ensure selection of words from the intended language.

Current models of CSI

As mentioned in the Introduction, there are two basic accounts regarding the dynamics of speech production that leads to CSI: the account of Howard et al. (2006), according to which lexical competition is indispensable for the CSI effect, and the alternative account introduced by Oppenheim et al. (2010), according to which such competition may exist but is not necessary for the CSI effect. Let us start by recalling how Howard et al. (2006) conceive lexical access and CSI. According to this model, concepts spread activation to related concepts and words which compete as candidates for lexical selection. Such competition is implemented through lateral inhibition, which entails that the higher the activation of a competitor, the longer it takes to select a target. When lexical selection occurs, the target representation (or the link from the concept to the target lexical representation) is persistently strengthened, resulting in an increased accessibility for future retrieval (repetition priming) of the same word but also in that word acting as a stronger competitor when retrieving related words later on. In this way, successive retrieval of semantically related words results in a cumulative increase in RTs.

An alternative to these dynamics was introduced by Oppenheim et al. (2010). In this model, the speed of lexical retrieval depends on the strength of the lexico-semantic links and this strength is adjusted with each new relevant speech experience after lexical selection takes place, both by strengthening the links of targets and by weakening those of competitors (e.g., error-based learning). In

Generalitat de Catalunya
Departament d'Economia
i Coneixement

this way, and in contrast to Howard et al. (2006), CSI can be implemented both assuming and dispensing with lexical competition. Let us briefly explain both possibilities.

If strengthening the links of targets is assumed as the mainly responsible mechanism for CSI, lexical competition can be maintained. Spreading activation results in words competing for selection as in Howard et al. (2006), but in contrast to the lateral inhibition assumed there, a lexical booster is in charge of guiding activation towards target words. By boosting activation of targets and competitors by the same constant (or antiboosting –inhibiting- competitors) in repeated processing cycles until a certain differential threshold is reached, lexical selection of the intended word is ensured even in situations of high competition. After selection, all activity returns to resting level, but the lexico-semantic links are adjusted according to the discrepancy between the desired activation of output nodes, and their actual activation before boosting (i.e., error-based learning), strengthening targets for future retrieval. In other words, naming “cat” will lead to an increased strength of the link from the conceptual information of “cat” (e.g., furry, animal, miaows) to the lexical entry (cat). Thus, when “dog” is the target on a subsequent trial, cat will receive more activation than normal through the features that were strengthened previously and are shared with dog (e.g., “animal” and “furry”), leading to a smaller initial activation difference between “cat” and “dog”. This means that more processing cycles will be required for the lexical booster to create a sufficient imbalance in activation between the target and competitors, which leads to increased RTs. Successive repetition of this process also leads to a cumulative cost in RTs.

Now let us turn to the version of Oppenheim et al. (2010) without lexical competition. In the simulation we are referring to, lexical selection is achieved when the target reaches an absolute threshold, independently of the activation levels of related words. CSI is still obtained because the links of competitors are weakened after lexical selection by the error-based learning mechanism. That is, after having named “dog”, the links to all other animals such as “horse” and “cat” are weakened. When “cat” has to be named later on, naming latencies will increase due to the weakened link from the conceptual information of “cat” shared with “dog" to the lexical representation “cat”. At the same time, after the eventual selection of “cat”, the links to all other animals which did not form part of the response such as “horse” are weakened again. Thus, when “horse” has to be named later on, the links

Agència de Gestió d'Ajuts Universitaris i de Recerca AGAUR
UNIÓ EUROPEA Fons Social Europeu

from the concept to the word were weakened twice (first during the production of "dog" and then again during the production of "cat") making selection even slower, resulting in a cumulative cost.

Having introduced the rudiments of the theoretical accounts of CSI, let us now examine how these may combine with different assumptions of bilingual language control in order to predict our findings.

Emergency exits for current bilingual models

As we have already seen, the combination of lexical competition and the kind of inhibitory control that is implemented in the ICM leads to an erroneous prediction: namely that semantic competitor effects should be absent or reduced in contexts of language alternation. Later on we will examine how other kinds of inhibitory control in combination with competitive selection could accommodate our results. But first, let us consider whether global inhibition through language membership could be maintained to explain our findings in the non-competitive framework of CSI introduced by Oppenheim et al. (2010). Strictly speaking, it could. As discussed above, within the non-competitive version of Oppenheim and colleagues (2010), CSI occurs due to weakening of the lexico-semantic links of competitors and the mechanism responsible for these adjustments operates based on *initial* activation levels. Thus, assuming that global inhibition is applied after the initial spread of activation, it would not affect the CSI effect. Importantly, however, this type of global inhibition is quite different from that conceptualized in the ICM, given that it does not have any consequences on posterior production. If it would have, then one would expect the CSI in a language alternating context to be of a greater magnitude compared to the language invariant context. This is because in addition to the weakening exerted by the learning mechanism based on the initial activation levels, there would also be the inhibition applied previously to semantically related co-activated items with incorrect language tags, predicting an additive cost in the selection of the target word (e.g., an increase in naming latencies due to weaker links and lower activation levels of the representations). This is not in line with the data pattern of the CSI that we obtained. Thus, the notion of bilingual language control through global inhibition is only tenable if (1) lexical selection is non-competitive (in contrast to what is assumed by the ICM); and (2) the global inhibition is very short-lived (1 trial; also in contrast to the ICM).

Generalitat de Catalunya Departament d'Economia i Coneixement

Regarding the language specific selection hypothesis, our results show that to the extent that lexical selection is by competition, such competition occurs also across languages. It has been argued elsewhere that apparent cross-language interference effects can be explained as within language effects in disguise (e.g., Costa et al., 1999). For instance, for the target “cat”, the translation “gato” becomes active and, assuming that it is not inhibited in any way, it may have a slightly increased activation for some time. In that way, when the word “perro” [dog] has to be named, “gato” may act as a stronger competitor. However, such a solution would not apply to the effect observed in the present study because both Howard et al. (2006) and Oppenheim et al. (2010) make a clear functional distinction between words that are targeted for production and their competitors. That is, in the models that assume lexical competition, CSI comes about because words (or the links from concepts to words) that are intended for production (either selected targets as in Howard et al., 2006 or desired targets as in Oppenheim et al., 2010) are persistently strengthened. Thus, according to Howard et al. (2006), although it may well be that translation words have an increased activity for some time, such residual activation is not sufficient to induce CSI. According to Oppenheim et al. (2010), translation words would not be strengthened either as they are not intended for production. This means that in a framework where lexical candidates compete for selection, the claim that lexical selection is language specific is not compatible with our findings. However, the general notion of language-specific selection could be maintained within a non-competitive lexical selection framework as long as the ability of focusing on one language is functional only after the initial spread of activation and as long as this mechanism has no consequences for subsequent activation levels or connection strengths.

In sum, largely modified versions of the ICM and the language specific selection hypothesis could still be maintained at the cost of losing their initial motivation, namely competitive lexical selection. However, more simple solutions to the issue of bilingual language control have already been proposed in the context of non-competitive lexical selection and should be rejected before adopting such complex solutions. In fact, at first sight, a language control mechanism that operates over lexical representations seems to be a redundant feature in a system where lexical selection is a non-competitive process, or as stated by Finkbeiner and colleagues (2006), “the hard problem” might disappear. For example, if language membership encoding at the preverbal level is sufficient to ensure that the intended target word receives more activation than the competitors in a similar vein to other conceptual information, then there is no obvious need for additional mechanisms of language control within a non-competitive framework. However, implementing language control at a preverbal level within Oppenheim et al.'s non-competitive explanation of CSI leads to the incorrect prediction that CSI should be of a larger magnitude within than between languages. This is because semantic category members from the same language as the target would receive more activation from the semantic system than semantic category members from other languages. This would lead to a greater weakening of words from the same language than of words from a different language. Thus, the simple solution of placing language as a feature at the semantic layer is insufficient if lexical selection is not by competition. Another recurrent way of accounting for semantic interference effects and bilingual language control within a framework without competitive lexical selection has been by locating them at the level of response exclusion (e.g., Finkbeiner and Caramazza, 2006; Finkbeiner et al., 2006; Mahon et al., 2007; Janssen et al., 2008; Dhooge and Hartsuiker, 2010, 2011; see Hall, 2011 for a review). This is not an option in the current context, since even though the models differ on where in the processing stream the CSI effect is induced (i.e., simultaneously with or after lexical selection), they both assume that CSI has an impact on the speed with which lexical selection can take place rather than on some previous or posterior process. In other words, there is agreement that CSI is a lexico-semantic effect. Thus, it seems like our data demonstrate that the simple solution to bilinguals speakers' "hard" problem (Finkbeiner et al., 2006), is too simple. Instead, there is a need for an additional mechanism of language control at the lexical level in order to create a substantial imbalance in activation between the intended and unintended language.

Arrived at this point, rather than continuing the search for emergency exits for bilingual models of speech production, we believe it may be more productive to put the monolingual accounts for lexical access and CSI under the magnifying glass and examine whether they may take charge of bilingual language control through any of their independently motivated features. We believe that lateral inhibition in Howard et al.'s (2006) model and the lexical booster in Oppenheim et al.'s (2010) model are good candidates for assuming such role.

The easy way out offered by existing monolingual models

Lateral inhibition in charge of language control

Let us first consider how the lateral inhibition that causes and resolves lexico-semantic competition in Howard et al. (2006) could account for our results while ensuring that lexical representations from the intended language are selected. In this model, semantic competitors are thought to delay target selection because of their inhibitory links to the target. In this way, the more active the competitors become, the more inhibition they will exert on the target, subsequently slowing down production. Eventually, the greater activity of targets will result in a greater inhibition of competitors than vice versa, thus ensuring that the intended word is selected (for a detailed description on the functionality of lateral inhibition, see Rumelhart & McClelland, 1982). One specific way in which target selection may be ensured is by detecting that the competing candidates have non-shared features with the target (e.g., "cat" and "dog" share the features "animal" and "furry", but "dog" contains the feature "bark" that is not shared with "cat"). The non-shared feature would thus result in a greater amount of activation of targets and thus inhibition of their competitors. Turning to bilingualism, one can assume that language membership functions as any other semantic feature (for similar proposals see: e.g., Finkbeiner et al., 2006; La Heij, 2005; Runnqvist & Costa, in press). Thus, "gato" would be co-activated along with "cat" because of their shared features, but the additional language feature would ensure that "cat" would inhibit "gato" to a greater extent than vice versa, ensuring lexical selection in the intended language. In short, language competition would be an instance of semantic competition and thus both emerge and be resolved within the lexical system.

Now, let us see whether such model would predict a CSI effect of the same magnitude within and between languages. Assuming that competition from different words does not only count in isolation, but rather adds to a net amount, such amount would be equally increased by competitors from the target language and the non-target language. For example, imagine that a target gets active with 10, same language competitors with 2 and other language competitors with 1. The net amount of competition would thus be 3. If "cat" has been named previously strengthening its activation with 1 and the target is "dog", the net amount of competition will be 4 (3 from "cat" +1 from" gato"). If "gato" has been named previously strengthening its activation with 1 and the target is "dog", the net amount of competition will also be 4 (2 from "cat" and 2 from "gato"). Thus, the lateral inhibition of Howard et al. (2006) offers a straightforward mechanism to explain the similar magnitude of the CSI effect across languages as well as a unified account for how language control is achieved in bilingual speakers and how lexico-semantic competition emerges and is resolved more generally.

The lexical booster in charge of language control

Another explanation which can capture our findings within a competitive framework and where language is not specified conceptually, capitalizes on another independently motivated mechanism, namely the "lexical booster" proposed by Dell, Oppenheim and collaborators (a.k.a. "traffic cop", e.g., Dell, Oppenheim, & Kittredege, 2008; Oppenheim et al., 2010). As mentioned previously, the lexical booster operates by boosting activation of the target and the competitors by the same constant, or by anti-boosting –inhibiting- activation of the competitors. The general purpose of this mechanism is that of aiding lexical selection when there is a high level of competition. Although so far only two sources of such competition have been specified (i.e., paradigmatic and syntagmatic interference), one could easily imagine that language could be implemented as a constraint on which items are boosted to ensure lexical selection in the correct language (either by boosting items of the target language or by anti-boosting –inhibiting- items of the non-target language, or both).

Crucially, the effects of the booster on activation levels will not have any consequences for semantic effects. This is because (a) in Oppenheim et al.'s (2010) model CSI and other similar interference effects occur due to an error-based learning mechanism that takes into account the *initial* levels of activation (i.e., *before* the booster mechanism comes in to play) to adjust the connection weights between semantic features and lexical entries; and (b) all items are assumed to return to their resting level of activation at the end of a given trial (i.e., there is no residual activation) which entails that whatever the booster does, it will not have any consequences for subsequent trials. This entails that as long as the initial activation of lexical items is blind to language, the lexical booster as a means of language control is compatible with a CSI effect of the same magnitude within and between languages.

Importantly, the lexical booster could also be in charge of language control within a non-competitive framework of lexical selection such as the one depicted by Oppenheim et al. (2010). In such a scenario, although the booster would no longer be relevant as a means of resolving lexical competition, it would still have an independent motivation in the need of directing activation towards a particular syntactic category to avoid syntagmatic interference. That is, when producing sentences, the speaker may experiment interference from highly activated words that form part of the intended message (e.g., in "the bird flies" the verb may interfere with the noun). The function of the booster is to direct activation towards the correct grammatical category at a given stage in the sentence (e.g., first "determiner", second "noun" and "third "verb"). Similarly, it could have the function of directing activation towards the correct language after an initial spread of activation that would be blind to language.

Conclusion

We reported five experiments testing the CSI within and across languages. Taken together, the results show a clear pattern. In language invariant contexts CSI is present within both languages of a bilingual speaker to a similar extent. In language alternating contexts CSI is present between languages, and the magnitude of this effect is similar to that observed within languages. By capitalizing on the detailed computational accounts of the phenomenon of CSI in the monolingual literature and combining these with our findings, we obtained novel constraints for theorizing about bilingual language control. None of the bilingual speech production models described in the literature specifies a language control mechanism capable to account for the results. The notion of language control through global inhibition as specified in the ICM (e.g., Green, 1998) can only be maintained if lexical selection is non-competitive and this type of inhibition has no consequences for subsequent speech behaviour; similarly, the notion of language control through selectivity (e.g., Costa et al., 1999) can only function in a lexical system which is non-competitive across the board (and not solely across languages); finally, also the proposal that language control takes place at the semantic level (Finkbeiner et al., 2006), can only be preserved if lexical selection *is* competitive. Thus, highly modified versions regarding the theoretical constructs of language control specified in the literature can be maintained. However, to do so, we must drop critical assumptions of the lexical architecture underlying those bilingual models that motivated them to begin with. Therefore, we discussed two alternatives to the traditional notions on bilingual language control which capitalize on independently motivated mechanisms and integrate bilingualism into speaker general dynamics, namely language control via lateral inhibition in a competitive system or a lexical booster sensitive to language in a competitive or non-competitive system.

Acknowledgements

This work was supported by grants from the Spanish government (PSI2008-01191, Consolider Ingenio 2010 CSD2007-00012) and the Catalan government (Consolidado SGR 2009-1521), a predoctoral grant from the Catalan government (FI) to E.R., a predoctoral grant from the Spanish government (FPU) to K.S, as well as funding from the European Research Council under the European Community's Seventh Framework Program (FP7/2007-2013 Grant agreement n° 263575).

Appendix A.

Participants were recruited based on their language dominance: they could only participate if they considered themselves as dominant in L1, spoke in L1 with the mother since birth, and considered their use of L1 to be more frequent than that of L2 (for the accuracy of such subjective measures to determine language dominance see Gollan, Weissberger, Runnqvist, Montoya & Cera, in press). Additionally, differences between L1 and L2 in the self-assessed language proficiency supported this dominance.

The self-assessed proficiency scores of the participants of all the experiments are presented in Table A. The mean age and standard deviation are given in years. The proficiency scores were obtained through a questionnaire filled out by the participants after the experiment. The scores are on a 4-point scale, where 4 = native-speaker level; 3 = advanced level; 2 = medium level; and 1 = low level of proficiency. The self-assessment index represents the average and standard deviation of the participants' responses in four domains (speech comprehension, speech production, reading, and writing).

Table A. Participants' self-rated language proficiency

Experiment	Age	L1-proficiency	L2-proficiency

1	20.12 (1.159)	3.93 (0.257)	3.82 (0.246)
2	21.50 (3.842)	3.94 (0.225)	3.72 (0.414)
3	20.75 (2.856)	4 (0)	3.74 (0.322)
4	20.54 (3.117)	3.96 (0.110)	3.82 (0.259)
5	20.30 (2.105)	3.99 (0.045)	3.87 (0.175)

Figure captions

Figure 1.
Schematic representation of a sequence of trials in Experiment 1.

Figure 2.
Average response times by participant and condition in Experiment 1. Response times are broken by response language (L1 in black, L2 in gray). The dashed lines represent the raw data, i.e., response times averaged by participant per condition. The vertical segments represent 95% confidence intervals around condition means across participants. The solid lines represent model fits for the effect of ordinal position, i.e., cumulative semantic interference. The parallel fit lines indicate that the effect is similar in the two conditions that are depicted.

Figure 3.
Schematic representation of a sequence of trials in Experiments 2 and 3.

Figure 4.
Average response times by participant and condition in Experiments 2 (panel a) and 3 (panel b). Response times in Experiment 3 are broken by response language (L1 in black, L2 in gray), since these two conditions were blocked between participants The dashed lines represent the raw data, i.e., response times averaged by participant per condition. The vertical segments represent 95% confidence intervals around condition means across participants. The solid lines represent model fits for the effect of ordinal position, i.e., cumulative semantic interference. The parallel fit lines indicate that the effect is similar in the three conditions that are depicted.

Figure 5.
Schematic representation of a sequence of trials in Experiments 4 and 5.

Figure 6.
Average response times by participant and condition in Experiments 4 (panel a) and 5 (panel b). Response times in Experiment 5 are broken by response language (L1 in black, L2 in gray), since these two conditions were blocked between participants. The dashed lines represent the raw data, i.e., response times averaged by participant per condition. The vertical segments represent 95% confidence intervals around condition means across participants. The solid lines represent model fits for the effect of ordinal position, i.e., cumulative semantic interference. The parallel fit lines indicate that the effect is similar in the three conditions that are depicted.

Figure 7.
Estimated CSI effects per category in experiments 2 (black) and 3 (gray) sorted by proportion of cognate items within the category. Categories are represented with their first letter. The black and grey lines represent estimates of the linear relationship between the two magnitudes in the two experiments. The pattern is very similar across experiments (see text for details).

Agència de Gestió d'Ajuts Universitaris i de Recerca
AGAUR

UNIÓ EUROPEA
Fons Social Europeu

CHAPTER 2: MIGHT INHIBITORY LANGUAGE CONTROL EXTEND THE BILINGUAL DISADVANTAGE TO MEMORY RETRIEVAL? (Runnqvist, E. & Costa, A. (2012). Is retrieval-induced forgetting behind the bilingual disadvantage in speech production? *Bilingualism: Language and Cognition*, 15 (02), 365-377.).

Generalitat de Catalunya
Departament d'Economia
i Coneixement

Abstract

In a previous study Levy et al. (2007) found that naming pictures in L2 impaired subsequent recall of the L1 translation words. This was interpreted as evidence for a domain-general inhibitory mechanism (RIF) underlying first language attrition. Because this result is at odds with some previous findings and theoretical assumptions we wanted to assess its reliability and replicate the experiment with various groups. Participants were first shown drawings along with their labels in the non-dominant language. Afterwards, they named 75% of these drawings in their first language or in their non-dominant language. Finally, participants' memory of all L1 words was tested through the presentation of a rhyme-cue. Recall of L1 words was better after naming pictures in the non-dominant language compared to when the picture was not named at all. This result suggests that speaking a second language protects rather than harms the memory of our first language.

Keywords: Bilingualism, Language production, Retrieval-Induced Forgetting, Bilingual disadvantage, First language attrition

Introduction

There is ample evidence that speaking two languages has collateral disadvantages in language processing. Bilinguals, compared to monolinguals, appear to have a slower and less reliable lexical access processing as revealed by an increase in naming latencies, decreased verbal fluency performance and more frequent tip of the tongue states even in their dominant language (e.g. Gollan, Montoya & Werner, 2002; Gollan, Montoya, Fennema-Notestine & Morris, 2005; Gollan, Fennema-Notestine, Montoya & Jernigan, 2007; Bialystok, Craik, & Luk, 2008; Ivanova & Costa, 2008; Luo, Luk & Bialystok, 2010; Sandoval, Gollan, Ferreira & Salmon, 2010). These difficulties are more pronounced when speakers experience L1 attrition; a phenomenon that refers to the gradual deterioration of the mother tongue (e.g. Köpke & Schmid, 2004; Schmid, 2002; Seliger, 1991). Despite the various studies that have addressed both the origin of the bilingual disadvantage in lexical access and the origin of L1 attrition, their causes are still open to debate (e.g. Köpke, 2007; Sandoval et al., 2010; see the recent special issue on L1 attrition edited by Schmid, 2010).

A recent study by Levy et al. (Levy, Mc Veigh, Marful & Anderson, 2007) provided evidence for a domain general memory mechanism that could underlie both the bilingual disadvantage in lexical access and L1 attrition. This mechanism is called RIF (Retrieval-Induced Forgetting), and it is supposed to reveal the effects of inhibitory processes that suppress interfering memory traces during memory recall. The authors argued that retrieving words from an L2 involves the activation of the corresponding memory traces of L1 translations, causing a strong interference effect. Given this interference, the selection of the desired word in L2 may entail the inhibition of the translation word in L1. As a consequence, and on the additional assumption that the L1 suppression increases cumulatively over the number of times one retrieves a given word in L2, the more a bilingual speaker uses the L2 the more suppression is inflicted on the memory traces of the L1 representations. Levy et al. (2007) argued that the decrease in the availability of L1 representations as a consequence of L2 use would be at the basis of L1 attrition. As they put it "(….) Native-language words for ideas used most often in the foreign language are most vulnerable to forgetting" (Levy et al., 2007, p. 33). The goal of the present study was to assess: a) the theoretical implications of the mechanism proposed by Levy et al., and b) the reliability of the observations themselves.

The mechanism proposed by Levy et al. (2007) to account for L1 attrition can also be used to explain the less severe bilingual disadvantages in speech production. For example, the finding that bilinguals are slower than monolinguals in lexical access could be explained as a reflection of the bilinguals' need to resolve competition between the target word and its translation. In fact, several proposals have been put forward before along these lines (e.g. Bialystok et al., 2008). Furthermore, this explanation of L1 attrition is based on the same mechanism that has been adopted by various models of bilingual lexical access to explain how bilinguals avoid lexical intrusions from the non-target language during speech production (e.g. Green, 1986, 1998). In sum, the mechanism proposed by Levy et al. (2007) provides a unified account for L1 attrition, the bilingual disadvantage in lexical access and the dynamics of bilingual lexical selection, in which inhibition of competing representations would be at the core of all three phenomena.

The basic mechanism of retrieval-induced forgetting and its extension to L1 attrition

RIF is an account of forgetting based on the observation that the recall of a given piece of information can impair subsequent retrieval of related knowledge (e.g. Anderson, Bjork & Bjork, 2000). In a typical RIF experiment, subjects study a series of category-exemplar pairs (FRUIT-ORANGE, FRUIT-BANANA; DRINK-BOURBON). Later on, half of the pairs of half of the categories are practiced in a retrieval phase where subjects have to recall the exemplar names (FRUIT-O_____). In a final test, subjects' memory of all exemplar items is tested (FRUIT-O___, FRUIT-B___, DRINK-B___). As expected, a robust effect of facilitation is observed for retrieval practiced items in comparison to baseline items (items of unpracticed categories). More interestingly, subjects' recall of unpracticed items of practiced categories is worse than memory of baseline items. This is precisely what the RIF effect is, and it has been argued to arise in the following manner. When subjects retrieve items during the practice phase, other related exemplars also become activated interfering with the target. This interference is resolved through the suppression of the competing items which leads to an impaired recall in the final test (e.g. Anderson et al., 2000; but see Camp, Pecher & Schmidt, 2007). RIF is not only found for members of taxonomic categories, but also for a variety of other situations such as memory of visuo-spatial objects and personality traits (e.g. Ciranni & Shiamura, 1999; MacRae & MacLeod, 1999; MacLeod & MacRae, 2001). This fact has lead some authors to propose that RIF is a domain general mechanism that is operative whenever there is a need to resolve interference between competing memory traces (e.g. Levy & Anderson, 2002), a situation that bilinguals have to face whenever speaking in one of their languages.

Levy et al. (2007) extended this view to the case of bilingualism, testing the hypothesis that the retrieval of L2 words will hinder the subsequent retrieval of the corresponding L1 translations. Three different phases were included in their experiment. First, subjects studied pictures along with their L2 labels. Second, they were asked to name the pictures ten, five, one or zero (baseline items) times in L1 or L2 depending on a color cue. Third, they were presented with prompt words and asked to retrieve the L1 labels that rhymed (Experiments 1 and 2a) or were semantically related (Experiment 2b) with the prompt word and that had been presented in the study phase.

The third phase in which participants are asked to retrieve the words in L1 is the critical one. In particular, what is important is the extent to which the retrieval of L1 words is affected by the number of times the corresponding L2 translations have been produced. When the L1 words were elicited by semantically related word prompts, L2 repetition benefited the retrieval of the L1 words. That is, the more times a word was named in L2 the better the recall of the corresponding L1 translation. Thus, no RIF was reported in this task. However, the crucial result upon which Levy et al. (2007) based their conclusions refers to the rhyming task, where RIF across languages was present. In this task, the retrieval of L1 words was impaired as the naming in L2 increased, but only for less fluent bilinguals (see general discussion for the effect of proficiency). This result was interpreted as revealing that the more times one uses an L2 word, the harder it becomes to retrieve its L1 translation.

However, a closer look at the experimental evidence reported by Levy et al. (2007) reveals that the RIF effect across languages is not very robust and it is numerically rather small. First, against the authors' arguments, the retrieval of words in L1 was not different for baseline, 1 and 5 repetitions. If anything, naming in L2 once helped subsequent retrieval of L1 translations as compared to baseline. That is, contrary to the authors' conclusions, to retrieve L2 words 1 or 5 times does not hamper the subsequent retrieval of the corresponding L1 translations. The authors appear to neglect these two results and focus on just one observation, namely the significant difference in recall rates of those pictures that were named 10 times in L2 (34%) as compared to baseline items (41%). Numerically this is a rather small effect. To appreciate this, it is worthwhile translating the percentages of recall into natural frequencies. The distribution of items in the different conditions was uneven. In the baseline set (pictures that were never named) there were 10 words, while in the 10 repetition set there were only 5 words. Hence, in natural frequencies, the 41% recall rate in the baseline condition corresponds to an average retrieval of 2 words, and the 34% retrieval in the 10 repetition condition corresponds to 1.7 words. This seems like a very small effect to account for L1 attrition. Similar problems arise when assessing the robustness of the interference effect produced by L2 naming on L1 recall in the phonological test in Experiment 2a. In this experiment, there were no differences between naming the pictures in L2 once, five times, or not naming the pictures at all. However, again there was a small effect after naming the pictures in L2 10 times compared to not naming them at all: 66% of the words (3.3 words in natural frequencies) were recalled in the former condition and 72% (3.6 words) in the latter. Nevertheless, this complex pattern of results might be used to argue that there is inhibition of the L1 translations during L2 production, but in doing so one should give an explanation of why neither 1 nor 5 repetitions are enough to see such inhibitory component. In our view, the small magnitude of the inhibitory effect and the lack of a systematic presence cast some doubts on the reliability and robustness of the phenomenon.

Leaving aside the issues regarding the robustness and reliability of the study of Levy et al. (2007), their results are remarkable considering previous evidence from the RIF literature. RIF effects are not present across the board, but rather are eliminated or reversed under various different conditions. For example, when subjects are encouraged to interrelate exemplars by focusing on the common properties of items, or when the practiced and unpracticed items belong to the same subcategory (i.e. "hoofed animals"), retrieval-induced forgetting can be reduced or eliminated (e.g., Anderson & McCulloch, 1999; Bauml & Hartinger, 2002). Similarly, it has been found that increasing the similarity between practiced and unpracticed items can lead to a facilitated recall of the latter, and this facilitation can even generalize to related items that did not belong to the initially studied materials (i.e. retrieval-induced facilitation, e.g., Anderson, Green & McCulloch, 2000; Chan et al., 2006; Chan, 2010).That is, RIF does not seem to be found when there is high level of semantic similarity between practiced and unpracticed items. Given that translation words are thought to have a very large semantically overlap (e.g. Zeelenberg & Pecher, 2003), one would expect to obtain retrieval induced facilitation rather than interference. Hence, the boundaries regarding the presence of RIF are not obviously met in the context of bilingual language production.

Retrieval-induced forgetting as the basis of L1 attrition: Theoretical implications

The conclusions reached by Levy et al. (2007) also have important theoretical implications when one puts them in the context of the directionality of the speech production process. Recall that the RIF effect was present (if anything) when the task involved phonological cue retrieval, but not when it involved semantic cue retrieval. From this observation the authors conclude that the inhibition occurs at the level of phonological representations and not at the level of lexico-semantic representations. This is an interesting proposal, but in a way it compromises the interpretation of RIF as the origin of L1 attrition since speech production is semantically and not phonologically driven. That is, we retrieve words because we want to convey a meaning not because they rhyme. And, it appears from Levy et al.'s results that when bilinguals have to retrieve words from the semantic system the supposed inhibition suffered by the phonological representations in L1 is completely irrelevant for the outcome of the speech production system. Consequently, the effects on lexical search prompted by phonological cues may be irrelevant for understanding how lexical access from the semantic system works.

As seen above, the observations of Levy et al. (2007) might be relevant for constraining models not only related to L1 attrition and the bilingual disadvantage in lexical access, but also models of general bilingual language control. However, it is also apparent that the results are in part inconsistent with previous observations and theoretical proposals, and also that the effect reported is small and not very robust. Thus, it is fundamental to assess the reliability and generalizability of these results. This was the goal of the present experiment, in which we followed closely the experimental paradigm used by Levy et al. (2007) and tested a total of 141 bilinguals. To preview the results, we failed to observe any RIF in L1 associated with naming in a weaker or non-dominant language. Instead, naming pictures in the non-dominant language facilitated the retrieval from memory of the corresponding translations in L1 as compared with no naming whatsoever.

Experiment: Is the Retrieval Induced Forgetting effect across languages a reliable phenomenon?

In this experiment, we tested 141 participants belonging to three different groups: low proficient L2 learners (group 1), medium proficient L3 learners (group 2) and high proficient bilinguals (group 3) (from now on, we will refer to L1 as the dominant language and L2/L3 as the non-dominant language). We decided to include these three different groups to cover a relatively wide range of bilinguals. This is important given the lack of information on the type of participants tested in the study of Levy et al. (2007). Participants were first shown drawings along with their non-dominant language labels. Afterwards, they had to name a set of these drawings in the dominant or the non-dominant language according to a color cue. While some pictures of the first phase were presented 1, 5 or 10 times, others were not named at all. The language in which each picture was named remained the same throughout this phase. Finally, subjects' memory of the dominant language labels was tested through the presentation of a written rhyme-cue. The crucial conditions were 1) language of picture-naming (dominant and non-dominant) and 2) number of repetitions (0, 1, 5, or 10). The dependent measure was the percentage of correctly recalled dominant language labels in the final test.

The presence of RIF would be indexed by a worse performance in retrieval of words in the dominant language associated with an increase of repetitions in the non-dominant language. According to Levy et al.'s results, we should expect to find RIF for the group that is low-proficient in the non-dominant language, and perhaps also for the group of medium-proficient speakers, but not for the group of high proficient bilinguals. In other words, following the results of Levy et al., we expected an interaction between the RIF effect and non-dominant language proficiency.

Method

Participants. Three groups of participants were included in the experiment. Group 1 comprised 56 native speakers of Spanish with low proficiency in English from the universities of País Vasco and Murcia. The languages tested were Spanish (dominant) and English (non-dominant). Group 2 comprised 53 native Spanish speakers with high proficiency in Catalan and medium proficiency in English from two universities located in Barcelona (University of Barcelona and Pompeu Fabra University). The languages tested were Spanish (dominant) and English (non-dominant). Group 3 comprised 32 native Spanish speakers with high proficiency in Catalan from the University of Barcelona. The languages tested were Spanish (dominant) and Catalan (non-dominant). Participants were asked to fill out a questionnaire regarding their language history and proficiency (see appendix 1 for more details).

Material. Forty drawings with concrete and unambiguous names in all testing languages were selected (See appendix 2 for a complete list of materials). All of them had non-cognate names in all testing languages and differed in their final syllable in Spanish (e.g. "vela", "espelma", "candle" [candle]; "perro", "gos", "dog" [dog]). Twelve filler drawings were included all of which also had different final syllables in Spanish. For the final test, 40 Spanish nouns that rhymed with the 40 Spanish picture labels but not with those of the other testing language were selected (e.g. "vela"-"tela", "perro"-"cerro").

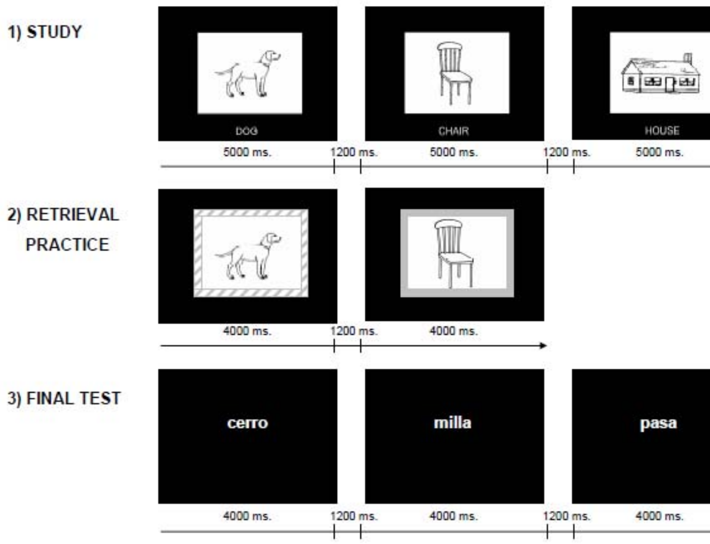
Procedure. There were three different phases in the Experiment (see figure 1). Each trial throughout the whole experiment started with a blank interval of 700 ms and a fixation point (an asterisk) presented for 500 ms. In phase 1 (study), all the 40 pictures were presented along with the corresponding word in the non-dominant language of the participant. The 40 pictures with the words were presented one at the time for 5 seconds and the task of the participants was to study the picture and its label. To control for effects of primacy and recency, twelve filler drawings were included (6 at the beginning and 6 at the end) so in total there were 52 trials in this phase of the experiment.

In the second phase (retrieval practice), participants were asked to name 75% of the pictures that had appeared in the first phase using the language indicated by a colored frame (e.g. green pictures named in the dominant language and red in the non-dominant language). Of this set of pictures, 25% were presented once (5 in green and 5 in red), 25% five times (5 in green and 5 in red), and 25% ten times (5 in green and 5 in red). In this way, the amount of trials in each of the two testing languages was equal. Each picture appeared with the same color frame throughout the experiment so that the language in which a given picture was named remained the same during this phase. The remaining 25% of pictures (10 pictures) did not appear in this phase and served as a baseline condition for the final test. To control for effects of primacy and recency, the same twelve filler drawings were included (6 at the beginning and 6 at the end), resulting in a total of 172 trials: ten pictures repeated once, ten repeated five times, ten repeated ten times and twelve fillers. Pictures were displayed until a response was given or for a maximum of 4 s. If participants did not answer, the correct name appeared on the computer screen during 500 ms.

In the third phase (final test), 40 words in the dominant language were presented, one at the time, on the computer screen. The task of the subjects was to provide a Spanish word that rhymed with the one presented on the computer screen and matched a previously viewed picture (*marco* - barco etc.). The rhyme-word was presented for a maximum of 4 s. and it disappeared with the detection of a response.

Figure 1.

Procedure of the cross-language RIF experiment



To control for effects of frequency, length and other possible confounding factors we decided to rotate the words through the experimental conditions. We thus created eight experimental lists because in the second phase of the experiment each picture could be named either in the dominant or in the non-dominant language and it could be named 0, 1, 5 or 10 times (2 languages*4 repetitions).

Data Analyses. Errors and omissions from the picture naming phase were analyzed by language. In the final rhyme-cued test the dependent measure was the percentage of correctly retrieved dominant language words. The analyses of the three groups of participants in the final test phase were firstly conducted separately. Two variables were included in a 2X4 ANOVA (language of picture naming: dominant vs. non-dominant; and number of repetitions (0, 1, 5 and 10).

Subsequently and given that the predictions are rather different for the dominant (a positive influence of repetition in subsequent word-retrieval in the dominant language) and the non-dominant language (a negative influence of repetition in subsequent word-retrieval in the dominant language), we also analyzed the results for the two languages separately. Finally, we conducted an omnibus analysis including the three groups of participants.

Only items that had been correctly named 80% of the time during the picture-naming phase are reported in the analysis in order to ensure that any effects could be unequivocally attributed to previous naming. Levy et al. (2007) did not specify whether they proceeded similarly (e.g., if they excluded or not the items that were not named or that were named erroneously). However, none of the effects reported below changed when all the responses were included in the analyses. Responses in which participants came up with a non-rhyming word or with a rhyming word that was not presented in the study phase were discarded from the analysis.

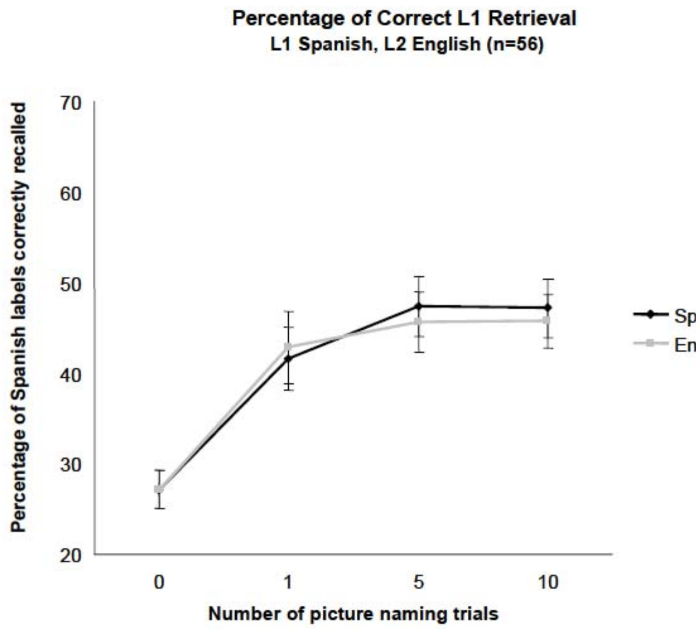
Results

Group 1 (Spanish dominant- English non-dominant).

Picture naming. Participants committed significantly more omissions in L2 (7%) than in L1 (1%) ($\underline{F}(1,55) = 56.353$, $\underline{MSE} = 16.607$, $\underline{p} = < .001$), but significantly more errors in L1 (8%) than in L2 (3%) ($\underline{F}(1,55) = 24.637$, $\underline{MSE} = 29.695$, $\underline{p} = < .001$).

Final test. The main effect of repetition was significant ($\underline{F}(3,165) = 19.96$, $\underline{MSE} = .047$, $\underline{p} = < .001$). Neither the main effect of language ($\underline{p} > .8$) nor the interaction between language and repetition was significant (p>.9). When analyzing the two languages separately, for the dominant language there was a main effect of repetition ($\underline{F}$ $(3,165) = 12.315$, $\underline{MSE} = .041$, $\underline{p} = < .001$), revealing a lower recall for the baseline condition (no repetitions) than for all the other conditions (all p's < .01). Thus, not surprisingly, repeating words in the dominant language helped subsequent access to those same words in the rhyming task. More important, for the non-dominant language, there was also a positive effect of repetition on recall performance ($\underline{F}$ $(3,165) = 8.832$, $\underline{MSE} = .050$, $\underline{p} = < .001$). This effect was significant for all repetition conditions against baseline (all p's < .01). That is, unlike in Levy et al.'s study (2007), <u>naming a picture in the non-dominant language facilitated later recall</u> of the translation word in the dominant language. Thus as is apparent in Figure 2, there was no RIF across languages.

Figure 2.

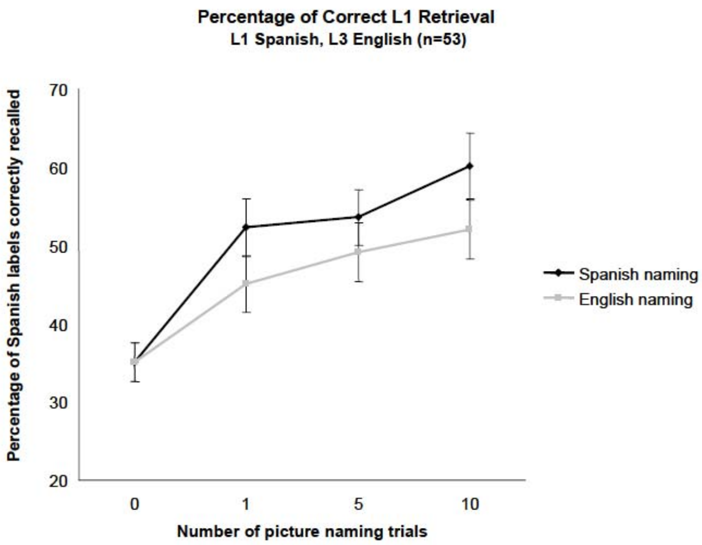


Group 2 (Spanish dominant-English non-dominant).

Picture naming. Participants committed significantly more omissions in L2 (3%) than in L1 (0.5%) (F(1,52) = 24.851, MSE = 5.581, p = < .001), but significantly more errors in L1 (8%) than in L2 (3%) (F(1,52) = 20.052, MSE = 34.936, p = < .001).

Final test. The main effects of repetition (F(3,156) = 13.208, MSE = .064, p = < .001) and language (F(1,52) = 5.236, MSE = .046, p =.026) were significant. The interaction between these variables was not significant (p>.5). When analyzing the two languages separately, for the dominant language there was a main effect of repetition (F (3,156) = 11.218, MSE = .052, p = < .001), revealing a lower recall for the baseline condition (no repetitions) than for all the other conditions (all p's < .01). Thus, again repeating words in the dominant language helped subsequent access to those same words in the rhyming task. More important, for the non-dominant language, there was also a positive effect of repetition on recall performance (F(3,156) = 5.272, MSE = .055, p = .002), with a significant difference between all repetition conditions against baseline (all p's < .02). Thus, no RIF across languages was present (see figure 3).

Figure 3.

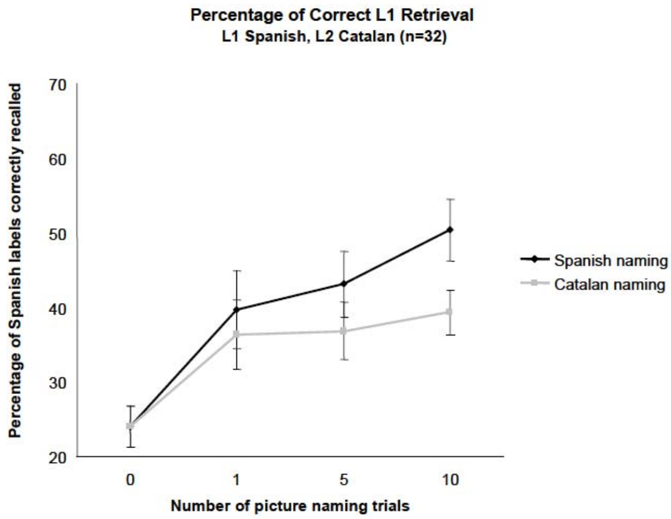


Group 3 (Spanish dominant- Catalan non-dominant).

Picture naming. Participants committed significantly more omissions in L1 (2%) than in L2 (1%) (F(1,31) = 4.030, MSE = 1.963, p = .053) and more errors in L1 (7%) than in L2 (4%) (F(1,31) = 7.599, MSE = 23.762, p = .01).

Final test. The main effect of repetition was significant (F(3,93) = 9.891, MSE = .052, p = < .001). The main effect of language approached significant values (F(1,31) = 2.888, MSE = .059, p =.099). No interaction between the two variables was observed (p>.4).When analyzing the two languages separately, for the dominant language there was a main effect of repetition (F(3,93) = 8.045, MSE = .049, p = < .001), revealing a better recall in all repetition conditions as compared to baseline (all p's < .01). For the non-dominant language, there was also a positive effect of repetition on recall performance: (F(3,93) = 3.523, MSE = .043, p = .025), an effect that was present for all repetition conditions when compared to baseline (all p's <.05). Thus, no RIF across languages was present (see figure 4).

Figure 4.



Omnibus analysis: A 4x2x3 ANOVA with "repetition" (0, 1, 5 and 10) and "language" (dominant/non dominant) as independent variables and group (1, 2, and 3) as a between subjects variable showed a significant main effect of repetition ($\underline{F}(3,414) = 38.981$, $\underline{MSE} = .055$, $\underline{p} = < .001$), language ($\underline{F}(1,138) = 6.188$, $\underline{MSE} = .051$, $\underline{p} =.014$) and group of participants ($\underline{F}(2,138) = 9.604$, $\underline{MSE} = .028$, $\underline{p} < .001$) but no significant interactions between any of these variables (all p's>.3).

Discussion

The results of this experiment showed a consistent pattern across the three groups of participants. Naming pictures in the dominant language led to a better recall of those very same words in a subsequent rhyming task, replicating previous observations (Levy et al., 2007). More important for our purposes is the effect of naming in the non-dominant language on the subsequent retrieval of the corresponding L1 translations in the rhyming task. We failed to see any detrimental effect of non-dominant language naming on dominant language recall. Instead, retrieval of dominant language words was substantially better when the corresponding translations were named in the non-dominant language than when they were not named at all.

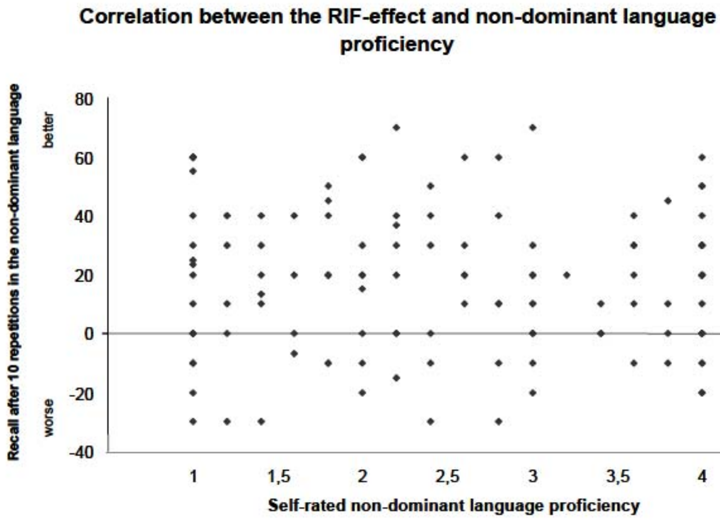
Furthermore, the performance in the rhyming task of the three groups of participants was rather similar. That is, the pattern of performance was independent of the participant's fluency in the non-dominant language and of whether they performed the task in their L2 or L3. Note however, that there was a main effect of group of participants in the omnibus analysis. That is, the group of medium proficient participants performed slightly better in the recall phase both in their L1 and L2 when compared to the other two groups. The origin of this difference is not clear to us, but importantly this effect did not interact with either the language of testing or the amount of repetitions. Recall that, aside from naming pictures in the dominant language, participants in group 1 and 2 also named pictures in English in which they were low and medium proficient respectively (probably the groups of participants that were most comparable to those of Levy et al., 2007), and participants in group 3 also named pictures in Catalan in which they were high proficient. Crucially, in none of the groups and in none of the three repetition conditions in the non dominant language did participants perform worse than in the baseline condition (the signature effect of RIF).

In fact, when considering the difference between the 10 repetitions condition in the non-dominant language and the baseline, only 28 participants out of 141 showed lower recall in the former condition (Figure 5 reports the collapsed analyses). A closer look at the individual data supports the notion that proficiency in the non-dominant language did not affect the percentage of recall in the rhyming task. As a proxy of language proficiency we took the average of the four questions answered by the participants regarding their knowledge of the non-dominant language. Also, we took the difference between the 10 repetitions condition and the baseline as an index of RIF (The data of the self-rated language proficiency for 18 participants was lost due to experimenter error. Therefore, the correlation only includes 123 participants). The correlation between these two indexes was near 0 ($r=.06$), and when the high proficient group was excluded the correlation was even worse ($r=.04$).Hence, it appears that proficiency does not predict recall performance in the dominant language after having named pictures in the non-dominant language.

In order to rule out the possibility that participants in our study employed a semantic strategy which yields facilitation as shown by Levy et al. (2007), we had a closer look at the errors elicited in the final test of the experiment (approximately 8% of the total amount of responses). The errors were divided into four types: (a) phonologically related extra-experimental items (those errors that rhymed with the target but did not form part of the experiment such as "raza" in response to "plaza"); (b) phonologically related experimental items (items that did form part of the experiment but only overlapped in the last two vowels with the rhyming cue such as "cara" in response to "plaza"); (c) phonologically unrelated extra-experimental items; and (d) phonologically unrelated experimental items. We observed that the amount of extra-experimental items was larger than that of experimental items. If subjects would have employed a semantic strategy or a memory driven strategy to perform the task, one would have expected to find a larger amount of experimental items than extra-experimental items in the errors. However, this was not the case. We also observed that the amount of phonologically related errors was larger than that of unrelated errors, which clearly supports that participants were employing a phonological strategy. We further compared the error pattern of those participants who showed RIF with that of the rest of participants and found no significant difference.

Before moving on, it is worth discussing a potential difference in the performance of the groups we tested. For two of the groups we observed the expected effect of better recall when the naming was performed in the same language as that of recollection, but for the low proficient bilinguals this difference was not significant. Note however, that in the omnibus analyses there was no interaction between language and group of participants. Still, one may argue that the low proficient bilinguals made use of a translation strategy during the picture naming phase (e.g. Kroll & Stewart, 1994). Independently of the use or not of this strategy, the direction of the effect we obtain for all groups is the same. While strategy differences due to proficiency may play a role in the magnitude of the difference in recall between languages, it does not seem to have an influence on the effect in general. In sum, the lack of a difference in dominant language recall after naming in the dominant and the non-dominant language does not compromise at all the fact that non-dominant language naming helps the retrieval of dominant language representations. If anything, it reaffirms the notion that words that are frequently used in the non-dominant language do not cause long-lasting retrieval impairment of their translations in the dominant language.

Figure 5.



Finally, we will dedicate some words to the apparently contradictory fact that participants in all three groups committed more errors in their dominant than in their non-dominant language. Recall that in the first phase of the experiment participants were familiarized with the picture labels in the non-dominant language, but not in the dominant language. Thus, participants knew which word they should use for each picture in the non-dominant language while they never received such an instruction for their dominant language, opening up the possibility of the use of alternative names for the pictures. Given the phonological nature of the final test of the experiment, such alternative names could not be considered as correct answers, thus increasing the error rate for the dominant but not for the non-dominant language.

General Discussion

The goal of this study was to explore the presence of retrieval induced forgetting (RIF) mechanisms in bilingual speech production. To do so, we assessed the reliability and generalizability of the RIF effect across languages reported by Levy et al. (2007). Recall that in their study there was an inhibitory effect of L2 speech production on L1 word recall. That is, those L1 words whose translations in L2 were produced several times were subsequently retrieved worse than those words that were not produced at all. To this end, we conducted a RIF experiment with 141 participants with differing proficiency levels in the non-dominant language. Two main findings were obtained.

a) Increasing the number of times a word is named in the dominant language increases the successful retrieval of this word in the same language in a subsequent rhyming task.

b) Increasing the number of times a word is named in the non-dominant language also increases the successful retrieval of its translation in the dominant language in a subsequent rhyming task.

While the first result is in line with the results of Levy et al. (2007), the second one is in clear conflict. At present we do not have an explanation for these contrasting results, that is, the presence vs. absence of RIF across languages in Levy's et al.'s (2007) and our study. However, we will now provide a general picture of the whole set of results that might be helpful in making sense of this discrepancy.

As mentioned in the introduction, the RIF effect across languages reported by Levy et al. was circumscribed. First, it was only present for those bilinguals with low fluency in their L2 (as indexed by differences in naming latencies between languages). For the other half of the participants no RIF whatsoever was observed. Second, even when assessing this sample, the RIF effect was present only between baseline and 10 repetitions, but not between baseline and 1 and 5 repetitions. Thus, the RIF effect across languages was only observed for half of the participants, and only comparing 10 repetitions vs. baseline. A closer look at our data reveals that only 28 participants showed a RIF effect, but none of the groups as a whole showed such an effect. All in all, this shows that, at best, the RIF effect across languages is a rather elusive phenomenon, and that caution needs to be exercised when deriving strong conclusions from Levy et al.'s observations.

Still, one may think of differences between the proficiency in the non-dominant language between the participants tested in our study and those of Levy et al.'s study as a potential source of the discrepant results. That is, it might be argued that the knowledge of the non-dominant language of our participants was greater than those of the previous study, and as a consequence we failed to observe any RIF effect across languages. This is a difficult argument to put forward since we do not know the characteristics of the sample used by Levy et al. Unfortunately, little information is given about the language history of the participants in Levy et al.'s study. We only know that they were undergraduate students who had recently completed at least one year of college-level Spanish –the non-dominant language-(we do not even know if English was the first language or whether they knew another language or not). Hence it is unclear how homogeneous the sample was. Some hint about the heterogeneity in the sample can be found in the post-hoc analyses conducted in their study, where the sample of 64 participants was split into two groups of 32 individuals each. The individuals were assigned to the two groups according to the difference in the speed with which they named the pictures in the dominant vs. the non-dominant language. This difference was taken as a proxy for differences in language proficiency. For one of the groups, the difference in naming latencies was quite large (more than 200 ms) being faster in English. However, for the other group naming latencies were faster in Spanish rather than in English. Thus, it appears that the sample was quite heterogeneous in Levy et al.'s study. Acknowledging that it is always difficult to control for cultural differences in learning an L2 between Europe and the USA, we believe that the wide spectrum of bilinguals tested in our study presumably guarantee that we covered the values of Levy et al.

Our data question the utility of RIF in language attrition. Even if one is tempted to conclude that RIF is present when the knowledge of the non-dominant language is very low, it is hard to imagine how this could be the cause of L1 attrition. This is because in real life, language attrition actually happens when people practice the non-dominant language enough. As the authors put it: "This phonological RIF arises precisely because **frequent use** (emphasis added) engages inhibitory control to achieve the fluency desired by foreign-language speakers" (Levy et al., 2007, p. 33). That is, frequent use is needed to promote RIF and, presumably, first language attrition. Consequently, one should expect RIF to be present also in the individuals tested in our experiment after a high amount of repetitions, or alternatively an interaction between language proficiency and RIF in the opposite direction to the one observed by Levy et al. (arguably, the more proficient speakers also use the non-dominant language more often than the less proficient ones). The fact that the vast majority of participants did not show the RIF effect and that there was no correlation between language proficiency and RIF seriously compromises the idea of this mechanism as the cause of first language attrition.

Note that the lack of a RIF effect across languages does not necessarily preclude the possibility that language control mechanisms make use of inhibitory processes (e.g. Kroll, Bobb, Misra, & Guo, 2008). Our results, however, do actually put constraints on whether one can use this inhibitory mechanism (at least as it is indexed by RIF) to account for the long lasting consequences of using a second language on the first language. In fact, there is another result that compromises the conclusion reached by these authors: "Native-language words for ideas used most often in the foreign language are most vulnerable to forgetting" (Levy et al., 2007, p. 33). This conclusion would predict that high frequency words should be more affected by L1 attrition than low frequency words. This is because high frequency words are by definition used more often than low frequency words. And, given that word frequencies tend to correlate across languages (at least in similar cultures), it is reasonable to expect that high-frequency words should suffer more RIF than low-frequency words. In this scenario, we would expect a reduced frequency effect in bilingual speakers when using their L1 as compared to monolingual speakers. However, recent results by Gollan, Montoya, Cera and Sandoval (2008) actually revealed the opposite, namely larger frequency effects for bilingual than for monolingual speakers. In fact, this effect was interpreted, against Levy et al.'s assumption, as revealing that low-frequency words are especially sensitive to bilingual disadvantages.

Before concluding, the observed facilitation between translation words fits well with some observations of RIF experiments. As advanced in the Introduction, when participants are presented with close semantic competitors (e.g., deer, moose,) practicing one of them (moose) helps rather than hinders the subsequent retrieval of the other item (deer) (e.g. Bauml & Hartinger, 2002; Anderson et al., 2000). In fact, crucially in our context, this facilitatory effect can generalize to other close semantically related items that have not been presented in the study phase (e.g., gazelle) (e.g. Chan et al., 2006). Close semantic overlap and generalization to unpractised items are two conditions met in our experiment, where translation words have a large semantic overlap, and practicing words in only one language exerts effects on the subsequent retrieval of its (unpractised) translation in the other language. That is, retrieving "cow" facilitated the recall of the Spanish translation "vaca" which overlaps largely in features with "cow" even though "vaca" had never been presented in the experiment. Interestingly, the contrastive effects of practicing items on the subsequent retrieval of semantically related items (sometimes facilitation and sometimes interference) have led to some authors to put forward explanations that might be useful to understand similar results in the bilingual field.

For the sake of simplicity we will only explain one of these models here, namely the feature suppression model by Anderson and Spellman (1995, see also Anderson et al., 2000), but it should be noticed that other similar proposals have been put forward in the literature (e.g. center-surround mechanisms, Carr & Dagenbach, 1990; Barnhardt, Glisky, Polster, & Elam, 1996; text processing account, Chan, 2009). In the feature suppression model, inhibition and activation are processes that are thought to co-occur whenever a given representation is retrieved from memory: related semantic representations have some shared and non-shared semantic features. When retrieving a given item, semantically related representations are activated by virtue of the shared features. However, the non-shared semantic features of semantically related items are inhibited, hence reducing the availability of such items. Thus, the presence of facilitation or inhibition would result from trade-off between the facilitation produced by shared features and the inhibition produced by non-shared features. For instance, retrieving "couch" would activate features such as "inanimate", "furniture" etc. shared with other representations such as "bed" and "table". At the same time, those non-shared features of "bed" and "table" such as "to sleep", "to put things on" etc. would be inhibited. In this way, the extent to and manner in which retrieving "couch" would subsequently affect the retrieval of "bed" depends on the shared non-shared features ratio: the bigger the ratio the more likely facilitation is observed. In this scenario, nothing but facilitation should be observed when retrieving words in language A that have been previously practiced in language B.

Interestingly, in the field of bilingual language production there is a similar literature about inhibitory and facilitatory effects of related words across languages. A closer look at all these phenomena reveals a striking parallelism to those observed in the memory literature. For example, several studies have reported facilitated or at least non-affected behaviour for translation-words (e.g. Francis, Augustini & Saenz, 2003; Costa et al., 1999; Costa & Caramazza, 1999; Lee & Williams, 2002). On the other hand, there are inhibitory effects in cross-language semantic competitor priming (Costa et al., 1999; Costa & Caramazza, 1999; Lee & Williams, 2002). Just to mention one example, naming a picture that is presented simultaneously with a written distracter in another language is easier if that word is the translation of the target to be produced, but harder if it is a semantic competitor, always compared to an unrelated condition (e.g. Costa et al., 1999; Costa & Caramazza, 1999). Thus, this contrast in on-line effects mimics the one between retrieval-induced facilitation and retrieval-induced forgetting and could be elegantly accounted for by the feature suppression model (but see Mahon, Costa, Peterson, Vargas & Caramazza, 2007 for an alternative account). Of course, such an account remains silent about how the bilingual speaker manages to restrict language production to only one language. It would be interesting to examine to what extent language membership behaves just as any other semantic feature in which case bilingual language production would be essentially similar to monolingual language production, only that one additional feature –language membership- would always have to be suppressed in order to produce speech in the intended language.

Conclusion

To conclude, the results of our study show that repeated production of words in a non-dominant language enhances memory for the translation words in the dominant language. Thus, although the mechanism of retrieval-induced forgetting might still be one of the many mechanisms that contribute to the bilingual disadvantage in speech production and to the phenomenon of first language attrition, the present results clearly calls into question such a unitary account.

*This research was supported by three grants from the Spanish Government (SEJ 2005/SEJ2005-01223/CONSOLIDER-INGENIO) and grants PSIC and PSI2008-00464/PSIC. Elin Runnqvist was supported by a Pre-doctoral fellowship from the Catalan Government (FI). The authors are grateful to Kristof Strijkers, and Dr. Clara Martin for their comments on previous versions of this manuscript.

Appendix 1. Language History and the Self-Assessed Proficiency for All Participants

The language history and self-assessed proficiency scores of the participants of all the groups are presented in Table A. The mean age and standard deviation are given in years. The “L2 /L3 onset” refers to the mean age (in years) at which participants started learning Catalan or English. The proficiency scores were obtained through a questionnaire filled out by the participants after the experiment. The scores are on a 4-point scale, where 4 = native-speaker level; 3 = advanced level; 2 = medium level; and 1 = low level of proficiency. The self-assessment index represents the average and standard deviation of the participants' responses in four domains (speech comprehension, speech production, reading, and writing).

Group	Age	Onset L2/3	Proficiency L1	Proficiency L2/3
1 (low proficient)	20.9 (2.5)	7.5 (2.8)	3.9 (.2)	1.6 (.6)
2 (medium proficient)	22.8 (3.9)	8.2 (2.5)	4.0 (.0)	2.6 (.6)
3 (high proficient)	21.4 (3)	3.3 (1.6)	4.0 (.0)	3.9 (.2)

CHAPTER 3. WILL THE BILINGUAL DISADVANTAGE EXTEND TO CONNECTED SPEECH AND WILL IT BE MODULATED BY FREQYENCY OF USE AND CROSS-LANGUAGE SIMILARITY? (Runnqvist, E., Gollan, T.H., Costa, A., & Ferreira, V.S. (under review). A frequency modulated disadvantage in bilingual sentence production. *Cognition*).

Abstract

Bilingual speakers access individual words less fluently, quickly, and accurately than monolinguals, particularly when accessing low-frequency words. Here we examined whether the bilingual speech production disadvantage would extend to full sentences and whether it would be modulated by structural frequency and syntactic properties of the bilingual speakers' other language. English monolinguals, Spanish-English bilinguals and Mandarin-English bilinguals were tested in a sentence production task conducted exclusively in English. Response times were modulated by bilingualism, structural frequency, and structural similarity across the bilingual speakers' two languages. These results refine our knowledge regarding the scope of the bilingual disadvantage and imply partially shared syntax across languages.

Keywords:

Sentence production; Bilingual disadvantage; Frequency effects; Cross-language interactivity

Introduction

Speaking two languages entails small but consistent disadvantages in linguistic processing compared to only speaking one language. For example, bilinguals take longer than monolinguals to name pictures and produce noun phrases (e.g., Gollan, Montoya, Cera, & Sandoval, 2008; Gollan, Slattery, Goldenberg, Van Assche, Duyck, & Rayner, 2011; Ivanova & Costa, 2008; Sadat, Martin, Alario, & Costa, 2011), even when speaking in their first and/or dominant language. This bilingual disadvantage is thought to be caused by a frequency lag in language usage with respect to monolingual speakers (e.g., Gollan et al., 2008; Gollan et al., 2011). Because bilinguals tend to use each of their two languages less frequently than monolinguals use their one language, linguistic information might be slightly less accessible for bilinguals (e.g., Gollan et al., 2008). A not mutually exclusive cause of the bilingual disadvantage might be the interference induced by simultaneously activated and competing representations from the language not in use (e.g., Kroll, Bobb, Misra, & Guo, 2008). Here we investigated whether bilingual disadvantages would extend to full sentence production and, if so, whether they would be modulated by structural frequency in the target language and syntactic similarity across the bilingual's two languages.

Our first goal was to assess whether the bilingual disadvantage would extend to production of full sentences. Evidence showing such a disadvantage for multi-word utterances such as "el coche rojo" [the red car] suggests that the bilingual disadvantage transcends single-word production (e.g., Sadat et al., 2011). Nevertheless, this conclusion might not apply to production of full sentences. Though noun-phrase production requires processes that are not shared with single-word production (e.g., word combination and gender agreement), it mostly involves retrieval of lexical information and lacks other processes needed to produce full sentences. Additionally, there are studies indicating that when a given structure is repeated throughout an entire experiment as in Sadat et al. (2011), production becomes more like picture naming, with speech-planning proceeding word by word (e.g., Martin, Crowther, Knight, Tamborello, & Yang, 2010). Normally, however, producing sentences involves additional processes (e.g., extended speech planning, verb inflection, etc.) compared to single-word production. To the extent that these additional processes are not specific to a particular language, they should decrease the bilingual disadvantage, because there would be no frequency lag or interference. In addition, bilinguals might develop more efficient processing for these operations, to compensate for their disadvantage in lexical retrieval, attenuating or eliminating the disadvantage in full sentence production.

Our second goal was to examine how structural frequency might modulate the bilingual disadvantage. Several studies indicate that the bilingual disadvantage is more pronounced for low-frequency words (e.g., Gollan et al., 2008; 2011; Ivanova & Costa, 2008). This finding led to the *frequency lag* hypothesis which assumes bilingual disadvantages arise from reduced use of each language relative to monolinguals and hence predicts that the disadvantage should be non-linear just as a frequency effect. Thus, assuming that statistical properties of syntax can have an impact on the onset of sentence production (in a similar way to structural complexity; e.g., Ferreira & Swets, 2002), it is possible that bilinguals may not show a big disadvantage for high frequency structures that are repeated massively every day, instead showing a bigger disadvantage for more low frequency structures with which they have less experience.

Our third goal was to seek for evidence of shared or interactive syntax modulating the bilingual disadvantage. If representations can be shared or interact across languages, then the bilingual disadvantage should be absent or reduced. This is presumably why bilinguals are not disadvantaged for producing proper names (Gollan, Bonanni & Montoya, 2005), and why cognates (translations with high semantic and phonological overlap) are named faster than non-cognates (e.g., Costa, Caramazza, & Sebastián-Gallés, 2000). Evidence showing transfer of syntactic properties across languages (e.g., Antón-Méndez, 2010) and cross-language syntactic priming (e.g., Bernolet, Hartsuiker, & Pickering, 2007; Hartsuiker, Pickering, & Veltkamp, 2004; Loebell & Bock, 2003; Shin & Christianson, 2009) suggests that syntax may be shared, or at least may interact during online processing, across languages. This might lead to full or partial frequency inheritance across languages for similar syntactic structures, attenuating, eliminating, or even inverting the bilingual disadvantage. Similarly, shared or interacting representations might lead bilinguals to experience less competition from the unintended language for structures that are similar across languages (i.e., Bates & MacWhinney, 1981; MacWhinney, 2005), which should also decrease the bilingual disadvantage. That is, bilinguals could be more or less disadvantaged in full sentence production as a function of the similarity of syntactic structures across languages.

To shed light on these possibilities, we tested bilinguals of different language combinations exclusively in the dominant language. The task induced production of sentences with structures that varied in English frequency and in cross-language similarity. Specifically, English monolinguals, Spanish-English bilinguals and Mandarin-English bilinguals produced English sentences by combining a verb ("push") or adjectival phrase ("is pink") with two nouns ("woman stroller") in the order they appeared on a computer screen ("the woman's stroller is pink") and response times were recorded. Target sentences involved two syntactic alternations: active versus passive voice (e.g., "the woman pushes the stroller" and "the stroller is pushed by the woman") and pre versus post modified possessive noun phrases (e.g., "the woman's stroller is pink" and "the stroller of the woman is pink"). Thus, by including four different types of syntactic constructions, all full sentences, we could be confident that participants engaged in the more complex speech planning and additional mental operations needed for full sentence production. Therefore, a modulation of bilingualism on response times would measure whether the disadvantage generalizes to the production of full sentences.

Within both syntactic alternations tested, one option is more frequent than the other in English (i.e., the active is more frequent than the passive, e.g., Bresnan, Dingare & Manning, 2001, and, at least for human possessors, pre-modified possessive NPs are more frequent than post-modified possessive NPs, e.g., Szmrecsanyi, 2009)[3]. This allowed us to assess whether response times would be affected by the statistical properties of the syntactic structures.

If bilinguals are disadvantaged more for low frequency representations in general, they should show larger frequency effects than monolinguals in both syntactic alternations. Furthermore, the frequency effects might be modulated by bilingual speakers' experience with a given structure through their other language (either directly by modulating the accessibility or indirectly by modulating the level of competition). In the case of the active-passive alternation, the English frequency distribution of the constructions is largely congruent with that of the other two languages (i.e., the active is more frequent than the passive in all three languages). However, in Spanish and in Mandarin, the frequency of the passive is even lower than in English (e.g., Blanco-Gómez, 2002; Xiao, McEnery, & Qian, 2006). Thus, if syntax interacts or is shared across languages, both Mandarin-English bilinguals and Spanish-English bilinguals might be relatively more disadvantaged in the passive construction compared to the active. Note that this prediction is identical to that of frequency-lag for bilinguals.

A more critical contrast for providing evidence for shared or interactive syntax was the pre versus post modified possessive NP alternation, where the frequency distribution differs across languages. In Mandarin, possessive NPs are always pre-modified (i.e., "女人的嬰兒車/Nu ren de ying er che" [woman possessive particle stroller]), while in Spanish all non-pronominal possessive NPs are post-modified (i.e., "la carreola de la mujer" [the stroller of the woman]). Thus, if syntax interacts or is shared across languages, Mandarin-English bilinguals might benefit from pre modified possessives being high frequency in Mandarin, decreasing the frequency lag with respect to English monolinguals for that construction. Similarly, they might be hampered by the lack of post modified possessive NPs in Mandarin, increasing the frequency lag for that construction. In either case, shared or interactive syntax

[3] Although we selected the structures based on their frequency differences, actives and passives might differ also in complexity or difficulty. Nonetheless, this is not true for the pre vs. post modified possessive alternation.

would predict an increased frequency effect for Mandarin-English bilinguals. Conversely, Spanish-English bilinguals might benefit from the post modified possessive NPs being high frequency in Spanish, or they might be hampered by the lack of pre modified possessive NPs in Spanish, or both, leading to a reduced frequency effect.

Method

Participants. 46 monolinguals, 50 early and high-proficiency Spanish-English bilinguals, and 49 early and high-proficiency Mandarin-English bilinguals took part in the experiment. Before the experiment, participants completed a language history and proficiency questionnaire. They also named pictures from the Multilingual Naming Test (Gollan, Weissberger, Runnqvist, Montoya & Cera, 2012) in English and (for bilingual participants) their other language. Finally, they completed a vocabulary and matrices reasoning tests (see Table 1). To test a homogeneous set of bilinguals, 10 Spanish-English bilinguals and 14 Mandarin-English bilinguals who reported being dominant in Spanish or Mandarin were excluded from the analyses. The remaining bilinguals were dominant in English as revealed by their naming accuracy. Participants without at least one observation in each condition were excluded (5 monolinguals, 6 Spanish-English bilinguals and 1 Mandarin-English bilingual), leaving 41 monolinguals, 34 Spanish-English bilinguals and 34 Mandarin-English bilinguals in the analyses[4].

Table 1. Description of participant characteristics. Numbers represent means of the indicated measure and numbers in parentheses represent standard deviations.

Group	*Age*	*Age first exposure to English*	*Age onset regular use of English*	*Age first exposure to other language*	*Age onset regular use of other language*	*Percent of daily use of English currently*	*Percent of daily use of English when growing up*
Monolinguals	20.29 (1.74)	0.30 (0.73)	1.63 (1.77)	10.39 (4.38)	13.00* (5.24)	98.95 (2)	95 (7)

[4] An additional analysis assessing frequency effects and the bilingual disadvantage across the two experimental blocks but including only a subset of participants with more observations per cell, elicited the same results for all main effects and interactions as the analysis reported in the article.

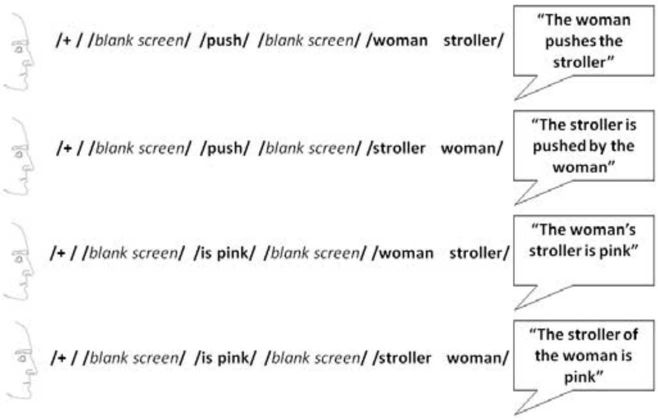
Mandarin bilinguals	20.21 2.38	3.13 (3.81)	4.97 (3.83)	0.51 (0.84)	2.65 (2.19)	90 (9)	66 (16)
Spanish bilinguals	21.44 (3.71)	2.95 (2.52)	4.81 (3.08)	0.51 (1.51)	1.56 (1.98)	86 (11)	65 (15)

Group	*Years of education*	*Scores matrices (1-46)*	*Scores shipley vocabulary (1-40)*	*Percent correct English picture naming*	*Percent correct other language naming*	*Self rated English proficency (1-10)*
Monolinguals	13.78 (1.42)	38.56 (4.24)	32.00 (2.85)	96 (3)	N.A.	9.58 (0.60)
Mandarin bilinguals	13.50 (0.99)	39.18 (4.64)	30.41 (3.16)	92 (5)	64 (21)	9.18 (0.80)
Spanish bilinguals	14.03 (1.40)	35.53 (4.08)	30.24 (3.09)	92 (4)	69 (15)	9.69 (0.50)

*the mean corresponds to the 12 monolingual participants who answered this question

Materials. The stimuli consisted of four agents (e.g., woman), eight objects (e.g., stroller), 32 verbs (e.g., push) and 32 adjectival phrases (e.g., is pink). The four agents were combined with eight objects each (e.g., woman, stroller). In turn, these combinations were paired with (a) a verb (e.g., woman, stroller, push), and (b) an adjectival phrase (e.g., woman, stroller, is pink). All combinations could elicit two different responses (e.g., "the woman pushes the stroller" or "the stroller is pushed by the woman" for the active vs. passive alternation, and "the woman's stroller is pink" and "the stroller of the woman is pink" for the pre vs. post modified NP alternation), resulting in a total of 128 possible sentences. Words were presented in black font (size 12) on a white background.

Figure 1.



Procedure. All trials (see Figure 1) began with a fixation cross (500 ms) and a blank screen (500 ms), followed by a first screen (1500 ms) containing either a verb (push) or an adjectival phrase (is pink). Then participants saw another blank screen (250 ms) and a second screen (until voice onset) containing an agent and an object ("woman stroller"). Participants were instructed to form grammatical English sentences as quickly as possible by combining the words from the first screen with those from the second screen, using the linear order in which the latter appeared ("woman stroller" or "stroller woman"). They were also instructed to add the determiner "the" where it was needed. Speakers were familiarized with this procedure in 12 practice trials before the experiment. The experiment consisted of two blocks of 32 sentences each (each speaker was presented with all combinations of agents, objects, verbs and adjectival phrases, but only had to produce one alternative of the syntactic alternation for a given combination). In this way, each speaker produced a total of 64 sentences (16 actives, passives, pre-modified possessive NPs and post-modified possessive NPs). Four different lists were created in which the combinations of agents, objects, verbs and adjectival phrases were counterbalanced, and the 32 combinations of each block were presented in a random order. The experiment was administered on Apple Macintosh computers running PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993). Vocal responses were recorded using a head-worn microphone connected to a PsyScope button box (measuring response times) and a standard recorder.

Analyses

Untargeted constructions (e.g., "the woman is pushing the stroller", "the stroller that the woman pushes" etc.), errors (partial repetitions, target word recall failures, omissions and voice-key triggering failures) and trials with response times three standard deviations below or above the average of each subject (2%) were removed. The remaining data were included in separate ANOVAs for each alternation (active/passive, pre/post modified possessive) with English structural frequency (low vs. high frequency) as a within-subjects variable and group (monolinguals vs. Mandarin-English bilinguals and monolinguals vs. Spanish-English bilinguals) as a between-subjects variable. Untargeted constructions and error rates were analyzed with analogous ANOVAs (see Tables 2, 3, and 4 and Figure 2.

Table 2. Response time ANOVAs (English structural frequency*group) for each alternation. Independent contrasts of English structural frequency and group are only shown when significant.

F1//F2	*PASSIVE VS. ACTIVE ALTERNATION*			*PRE VS. POST MODIFIED NP ALTERNATION*		
Contrast	*English structural frequency*	*Group*	*English structural frequency*Group*	*English structural frequency*	*Group*	*English structural frequency*Group*
Monolinguals, Mandarin bilinguals	F1(1, 73) =60.86, MSE=454009, p<.001, η_p^2=.455 // F2(1, 62) =122.444, MSE=212146, p<.001, η_p^2=.664	F1(1, 73) =9.29, MSE=2245843, p=.003, η_p^2=.113 // F2(1, 62) =68.194, MSE=156856, p<.001, η_p^2=.524	F1(1, 73) =5.41, MSE=454009, p=.023, η_p^2=.069 // F2(1, 62) =15.760, MSE=212146, p<.001, η_p^2=.203	F1(1, 73) =21.84, MSE=458141, p<.001, η_p^2=.230 // F2(1, 62) =60.544, MSE=116403, p<.001, η_p^2=.494	F1(1, 73) =3.74, MSE=1446850, p=.057, η_p^2=.049 // F2(1, 62) =17.497, MSE=129878, p<.001, η_p^2=.220	F1(1, 73) =3.96, MSE=458141, p=.050, η_p^2=.051 // F2(1, 62) =6.392, MSE=116403, p=.014, η_p^2=.093
Monolinguals, Spanish bilinguals	F1(1, 73) =72.52, MSE=376358, p<.001, η_p^2=.498 // F2(1, 62) =29.740, MSE=511005, p<.001, η_p^2=.324	F1(1, 73) =11.94, MSE=1948192, p=.001, η_p^2=.141 // F2(1, 62) =62.443, MSE=207656, p<.001, η_p^2=.502	F1(1, 73) =6.26, MSE=376358, p=.015, η_p^2=.079 // F2<1	F1(1, 73) =19.22, MSE=240423, p<.001, η_p^2=.208 // F2(1, 62) =17.784, MSE=187645, p<.001, η_p^2=.223	F1(1, 73) =8.34, MSE=1392159, p=.005, η_p^2=.102 // F2(1, 62) =39.254, MSE=180693, p<.001, η_p^2=.388	F1<1 // F2<1
Monolinguals	F1(1, 40) =25.88, MSE=290128, p<.001, η_p^2=.393 // F2(1, 31) =30.501, MSE=175092, p<.001, η_p^2=.496			F1(1, 40) =6.83, MSE=266310, p=.013, η_p^2=.146 // F2(1, 31) =22.616, MSE=71003, p<.001, η_p^2=.422		
Mandarin bilinguals	F1(1, 33) =32.62, MSE=652652, p<.001, η_p^2=.497 // F2(1, 31) =96.223, MSE=249201, p<.001, η_p^2=.756			F1(1, 33) =13.46, MSE=690663, p=.001, η_p^2=.290 // F2(1, 31) =38.230, MSE=161803, p<.001, η_p^2=.552		
Spanish bilinguals	F1(1, 33) =43.44, MSE=480880, p<.001, η_p^2=.568 // F2(1, 31) =12.107, MSE=846919, p=.002, η_p^2=.281					

Contrast	Active	Passive	Pre modified possessive	Post modified possessive
Monolinguals, Mandarin bilinguals	F1(1, 73) =6.162, MSE=730279, p=.015, ηp^2=.078 // F2(1, 62) =15.695, MSE=66249, p<.001, ηp^2=.202	F1(1, 73) =9.550, MSE=1969572, p<.003, ηp^2=.116 // F2(1, 62) =42.940, MSE=302753, p<.001, ηp^2=.409		F1(1, 73) =4.696, MSE=1436766, p=.033, ηp^2=.060 // F2(1, 62) =19.002, MSE=147803, p<.001, ηp^2=.235
Monolinguals, Spanish bilinguals	F1(1, 73) =10.455, MSE=517284, p=.002, ηp^2=.125 // F2(1, 62) =22.114, MSE=199542, p<.001, ηp^2=.263	F1(1, 73) =11.183, MSE=1807266, p=.001, ηp^2=.133 // F2(1, 62) =17.243, MSE=519119, p<.001, ηp^2=.218		

Table 3. Untargeted constructions ANOVAs (English structural frequency*group) for each alternation.

Independent contrasts of English structural frequency and group are only shown when significant.

F1/F2	*PASSIVE VS. ACTIVE ALTERNATION*			*PRE VS. POST MODIFIED NP ALTERNATION*		
Contrast	*English frequency*	*Group*	*English frequency*Group*	*English frequency*	*Group*	*English frequency*Group*
Monolinguals, Mandarin bilinguals	F1(1, 73) =22.332, MSE=426, p<.001, ηp^2=.234 //F2(1, 62) =57.636, MSE=134, p<.001, ηp^2=.482	F1(1, 73) =2.883, MSE=620, p=.094, ηp^2=.038 //F2(1, 62) =7.337, MSE=201, p=.009, ηp^2=.106	F1<1 //F2<1	F1<1 //F2<1	F1<1 //F2<1	F1<1 //F2(1, 62) =5.255, MSE=23, p=.025, ηp^2=.078
Monolinguals, Spanish bilinguals	F1(1, 73) =46.761, MSE=322, p<.001, ηp^2=.390 //F2(1, 62) =113.233, MSE=108, p<.001, ηp^2=.646	F1(1, 73) =1.814, MSE=591, p=.182, ηp^2=.024 //F2(1, 62) =4.370, MSE=227, p=.041, ηp^2=.066	F1(1, 73) =2.936, MSE=322, p=.091, ηp^2=.039 //F2(1, 62) =7.486, MSE=108, p=.008, ηp^2=.108	F1<1 //F2<1	F1<1 //F2<1	F1(1, 73) =5.557, MSE=94, p=.021, ηp^2=.071 //F2(1, 62) =12.059, MSE=32, p=.001, ηp^2=.163
Monolinguals	F1(1, 40) =14.577, MSE=320, p<.001, ηp^2=.267 // F2(1, 31) =38.859, MSE=87, p<.001, ηp^2=.556			F1(1, 40) =1.504, MSE=114, p=.227, ηp^2=.036 // F2(1, 31) =3.901, MSE=28, p=.057, ηp^2=.112		
Spanish bilinguals	F1(1, 33) =33.184, MSE=325, p<.001, ηp^2=.501 // F2(1, 31) =74.815, MSE=129, p<.001, ηp^2=.707			F1(1, 33) =5.202, MSE=69, p=.029, ηp^2=.136 //F2(1, 31) =8.347, MSE=36, p=.007, ηp^2=.212		

Contrast	*Active*	*Passive*	*Pre modified possessive*	*Post modified possessive*
Monolinguals, Mandarin bilinguals				F1<1 // F2(1, 62) =4.724, MSE=19, p=.034, ηp^2=.071
Monolinguals, Spanish bilinguals		F1(1, 73) =2.787, MSE=723, p=.099, ηp^2=.037 // F2(1, 62) =9.827, MSE=183, p=.003, ηp^2=.137		F1(1, 73) =2.300, MSE=144, p=.134, ηp^2=.031 // F2(1, 62) =12.573, MSE=23, p=.001, ηp^2=.169

Table 4. Error-rate ANOVAs (English structural frequency*group) for each alternation. Independent contrasts of English structural frequency and group are only shown when significant.

F1/F2	*PASSIVE VS. ACTIVE ALTERNATION*			*PRE VS. POST MODIFIED NP ALTERNATION*		
Contrast	*English frequency*	*Group*	*English frequency*Group*	*English frequency*	*Group*	*English frequency*Group*

Monolinguals, Mandarin bilinguals	F1(1, 73) =22.769, MSE=30, p<.001, ηp^2=.238 //F2(1, 62) =28.399, MSE=21, p<.001, ηp^2=.314	F1<1 //F2<1	F1<1 //F2(1, 62) =1.117, MSE=21, p=.295, ηp^2=.018	F1(1, 73) 1.960, MSE=30, p=.166, ηp^2=.026 //F2(1, 62) 1.564, MSE=34, p=.216, ηp^2=.025	F1<1 //F2<1	F1(1, 73) 4.322, MSE=30, p=.041, ηp^2=.056 //F2(1, 62) 3.3285, MSE=34, p=.075, ηp^2=.050
Monolinguals, Spanish bilinguals	F1(1, 73) =23.653, MSE=34, p<.001, ηp^2=.245 //F2(1, 62) =23.660, MSE=30, p<.001, ηp^2=.276	F1(1, 73) =3.025, MSE=65, p=.086, ηp^2=.040 //F2(1, 62) =6.447, MSE=27, p=.014, ηp^2=.094	F1<1 //F2<1	F1<1 //F2<1	F1<1 //F2<1	F1(1, 73) =2.021, MSE=32, p=.159, ηp^2=.027 //F2(1, 62) =1.865, MSE=27, p=.177, ηp^2=.029
Monolinguals	F1(1, 40) =18.180, MSE=29, p<.001, ηp^2=.312 // F2(1, 31) =34.023, MSE=12, p<.001, ηp^2=.523					
Spanish bilinguals	F1(1, 33) =7.613, MSE=40, p=.009, ηp^2=.187 // F2(1, 31) =6.264, MSE=48, p=.018, ηp^2=.168					
Mandarin bilinguals				F1(1, 33) =6.032, MSE=28, p=.019, ηp^2=.155 // F2(1, 31) =4.796, MSE=34, p=.036, ηp^2=.134		

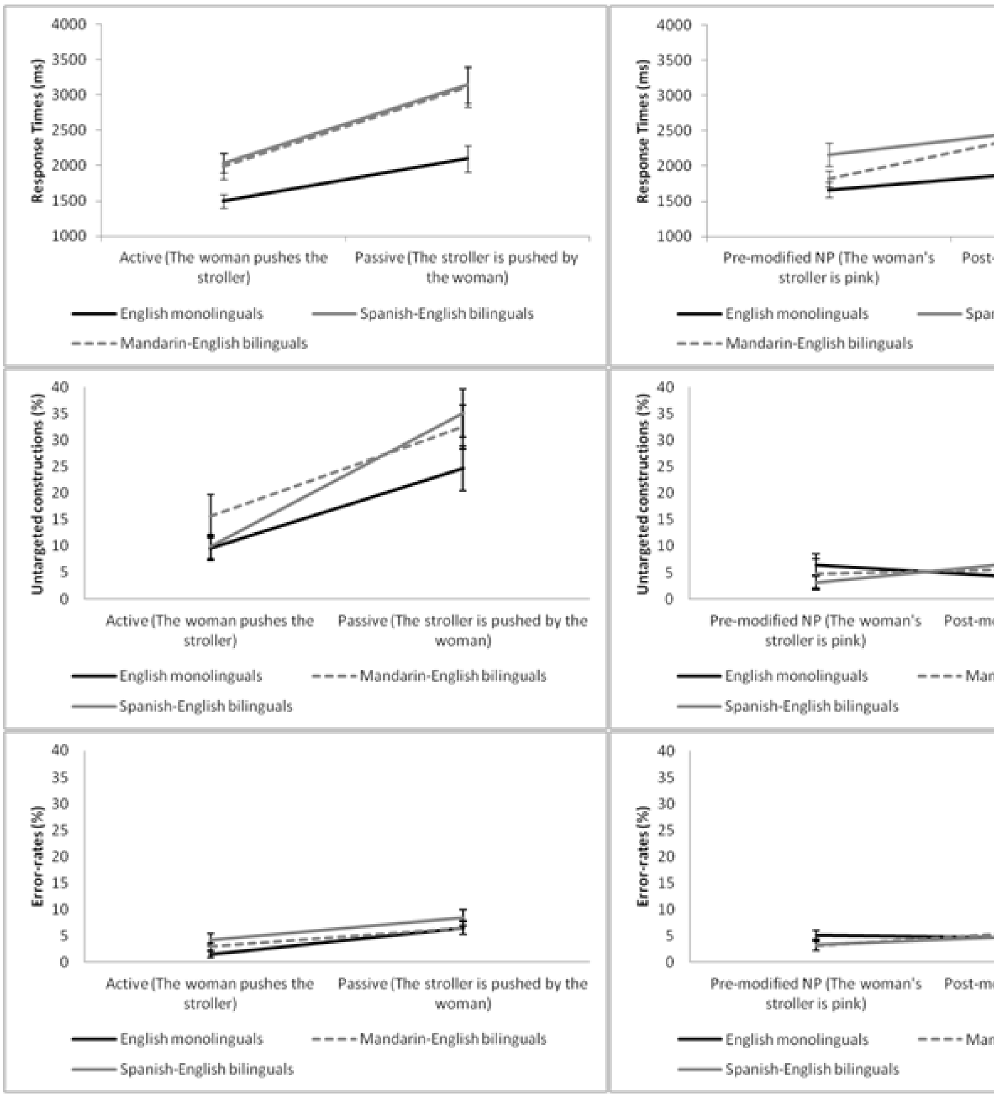
Contrast	*Active*	*Passive*	*Pre modified possessive*	*Post modified possessive*
Monolinguals, Spanish bilinguals	F1(1, 73) =4.978, MSE=27, p=.029, ηp^2=.064 // F2(1, 62) =9.429, MSE=13, p=.003, ηp^2=.132			

Results and Discussion

Bilingual disadvantage. Bilinguals were slower overall than monolinguals, revealing a bilingual disadvantage regardless of the combination of languages spoken. This was supported by main effects of group in both alternations and both group contrasts. Where such group effects interacted with English structural frequency, separate ANOVAs for each construction were conducted which corroborated the presence of a bilingual disadvantage in all constructions, except for the Mandarin-English bilinguals who were not slower than monolinguals for the pre modified possessives, just as predicted by the shared or interactive syntax hypothesis. Mandarin-English bilinguals produced more untargeted sentences than monolinguals in the active-passive alternation, and Spanish-English bilinguals produced more untargeted sentences than monolinguals only for the passives. The pre versus post modified possessive alternation showed a more complex pattern which we will refrain from interpreting given the low rate of untargeted responses in this alternation; the same holds for error-rates.

These findings extend the scope of bilingual disadvantages to full-sentence production, and demonstrate that bilinguals do not develop more efficient sentence production processes so as to compensate for their lexical access disadvantages. Though consistent with the finding of a bilingual disadvantage for production of noun-phrases (e.g., Sadat et al., 2011), the current data go beyond these data in several ways. The present study involved different types of larger and syntactically more complex utterances than those of Sadat et al. (2011). Thus, additional morpho-syntactic operations such as verb inflection and extended speech planning were required to perform the task. We hypothesized that bilinguals might compensate for their less efficient word retrieval by developing greater efficiency in these language non-specific processes, attenuating in this way the bilingual disadvantage. Our findings suggest that this is not the case.

Figure 2.



Structural frequency and other language modulations of the bilingual disadvantage. In the active-passive alternation, both Spanish-English and Mandarin-English bilinguals showed larger frequency effects than monolinguals. This was supported by interactions of English frequency by group for all comparisons between monolinguals and bilinguals (although for Spanish-English bilinguals this interaction only reached significance in the F1 analysis). This result might be caused both by a frequency-lag in English and by an influence of bilingual speakers' other language because passives are even lower frequency in Mandarin and Spanish than in English.

In the pre versus post modified NP alternation, Mandarin-English bilinguals showed larger frequency effects than monolinguals, as predicted by the shared or interactive syntax hypothesis, which was confirmed by an English frequency by group interaction. This interaction was at least partially driven by the lack of a disadvantage in the pre modified possessive construction that is high frequency in Mandarin. Also as predicted by the shared or interactive syntax hypothesis, the Spanish-English bilinguals did not show a larger frequency effect than monolinguals. That is, if only English structural frequency mattered (if syntax is not shared or interactive), a larger disadvantage for the low frequency construction for both bilingual groups should have been observed. In this respect the results largely fit the predictions of the shared-syntax hypothesis. One aspect that did not fit this account as well was that Spanish-English bilinguals did not show a *smaller* frequency effect than monolinguals as would have been predicted if syntax were completely shared, and frequency of syntactic structures fully inherited across languages. On this view, Spanish-English bilinguals should have been most disadvantaged of all structures for the pre-modified possessives that do not exist in Spanish, and should have produced the post-modified possessives more quickly than monolinguals because these are very high-frequency in Spanish and low frequency in English. Although we observed a disadvantage for the pre modified possessives, we did not observe a Spanish-English bilingual advantage for the post modified possessives. One possible reason why the influence of non-target language syntax was clearer in the English-Mandarin group is that in Mandarin, all possessive forms are pre-modified, while in Spanish only part of the possessive forms (the non-pronominal ones) are post-modified. Alternatively, Spanish-English bilinguals might not have been faster than monolinguals in the post modified possessives because the degree of syntactic interactivity or sharing across languages is restricted, and bilinguals must also keep track of the separate frequencies of syntactic structures in each language. Taken together, these results lend support to accounts that propose interactive or partially shared syntax across languages (e.g., Bates & MacWhinney, 1981; MacWhinney, 2005; Hartsuiker et al., 2004; Bernolet et al., 2007) and suggest that such interactivity is relevant for language production as it occurs for bilinguals in contexts in which only one language is used.

To conclude, like single word production, bilingual sentence production shows disadvantages relative to monolinguals, and like in word production these disadvantages are modulated both by frequency of use and by cross-language interactions.

Acknowledgements

This work was supported by grants from NICHD (HD050287 and HD051030), and the Spanish government (PSI2008-01191, Consolider Ingenio 2010 CSD2007-00012), and the Catalan government (Consolidado SGR 2009-1521), a predoctoral grant (FI) and a travelling grant (BE) from the Catalan government.

Figure legends

Figure 1. Schematic representation of the experimental procedure illustrated through an example of each type of target structure.

Figure 2. Response times (top), untargeted constructions (middle) and error rates (bottom) for the active and passive constructions (left) and the pre-modified possessive NP and post-modified possessive NP constructions (right) broken down by group of speakers (English monolinguals, English-Spanish bilinguals and English-Mandarin bilinguals). Error bars represent standard error.

References

Abdel Rahman, R., & Melinger, A. (2009). Semantic context effects in language production: A swinging lexical network proposal and a review. Language and Cognitive Processes, 24, 713–734.

Alario, F.-X., Segui, J., & Ferrand, L. (2000). Semantic and associative priming in picture naming. The Quarterly Journal of Experimental Psychology, 53 (A), 741–764.

Alario, F.-X., & Moscoso del Prado Martin, F. (2010). On the origin of the cumulative semantic inhibition effect. Memory & Cognition, 38, 57-66.

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. Psychonomic Bulletin & Review, 7, 522–530.

Anderson, M. C., & McCulloch, K. C. (1999). Integration as a general boundary condition on retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25, 608–629.

Anderson, M. C., Green, C., & McCulloch, K. C. (2000). Similarity and inhibition in long-term memory: Evidence for a two-factor model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1141–1159.

Anderson, M. C., & Spellman, B. A. (1995). On the status of inhibitory mechanisms in cognition: Memory retrieval as a model case. *Psychological Review*, 102, 68–100.

Antón-Méndez, I. (2010). Whose? L2-English speakers' possessive pronoun gender errors. Bilingualism: Language and Cognition, 1–14.
Baayen, R. H. (2008) Analyzing Linguistic Data. A Practical Introduction to Statistics Using R. Cambridge University Press.
Baayen, R.H., Davidson, D.J. & Bates, D.M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language, 59, 390-412.
Barnhardt, T. M., Glisky, E. L., Polster, M. R., & Elam, L. (1996). Inhibition of associates and activation of synonyms in the rare-word paradigm: Further evidence for a center-surround mechanism. *Memory & Cognition*, 24, 60–69.
Bates, E., & MacWhinney, B. (1981). Second language acquisition from a functionalist perspective: Pragmatic, semantic and perceptual strategies. In H. Winitz (Ed.), *Annals of the New York Academy of Sciences conference on native and foreign language acquisition*, New York: New York Academy of Sciences, 190-214.
Bauml, K., & Hartinger, A. (2002). On the role of item similarity in retrieval-induced forgetting. *Memory*, 10, 215–224.
Bernolet, S., R. J. Hartsuiker, and M. J. Pickering (2007). Shared syntactic representations in bilinguals: evidence for the role of word-order repetition. Journal of Experimental Psychology. Learning, Memory, and Cognition 33(5).931–49.
Bialystok, E., Craik, F., Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34 (4), 859–873.
Blanco-Gómez, M.L. (2002). Hiding the agent in English and Spanish newspaper articles: The periphrastic passive. In Marín Arrese, J.I. (ed.). *Conceptualization of events in newspaper discourse: mystification of events and degree of implication in news reports*. Universidad Complutense de Madrid
Bloem, I., & La Heij, W. (2003). Semantic facilitation and semantic interference in word translation: Implications for models of lexical access. Journal of Memory and Language, 48, 468–488.
Bloem, I., van den Boogaard, S., & La Heij, W. (2004). Semantic facilitation and semantic interference in language production: Further evidence for the conceptual selection model of lexical access. Journal of Memory and Language, 51, 307-323.
Box, G. E. P., & Cox, D. R. (1964). "An analysis of transformations". Journal of the Royal Statistical Society, Series B 26 (2), 211–252.
Bresnan, J., S. Dingare, & Manning ,C. D. (2001). Soft constraints mirror hard constraints: Voice and person in English and Lummi. In *Proceedings of the LFG 01 Conference*. CSLI Publications.
Brown, A. S. (1981). Inhibition in cued retrieval. Journal of Experimental Psychology: Human Learning and Memory, 7 (3), 204–215.
Camp, G., Pecher, D., & Schmidt, H.G. (2007) No retrieval-induced forgetting using item-specific independent cues: Evidence against a general inhibitory account. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(5), 950-958.
Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14, 177–208.
Carr, T. H., & Dagenbach, D. C. (1990). Semantic priming and repetition priming from masked words: Evidence for a center-surround attentional mechanism in perceptual recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition,* 16**,** 341–350.

Chan, J. C. K., McDermott, K. B., & Roediger, H. L., III (2006). Retrievalinduced facilitation: Initially nontested material can benefit from prior testing of related material. *Journal of Experimental Psychology: General*, 135, 533–571.
Chan, J. C. K. (2009). When does retrieval induce forgetting and when does it induce facilitation? Implications for retrieval inhibition, testing effect, and text processing. *Journal of Memory and Language*, 61, 153–170.
Christoffels, I. K., Firk, C., & Schiller, N. (2007). Bilingual language control: An event-related brain potential study. Brain Research, 1147, 192–208.
Ciranni, M.A. & Shimamura, A.P. (1999) Retrieval-induced forgetting in episodic memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 25, 1403–1414
Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphical system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, and Computers*, 25, 257-271.
Colomé, A. (2001). Lexical activation in bilinguals' speech production: language-specific or language-independent? *Journal of Memory and Language*, 45, 721-736.
Costa, A., & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish-English and English-Spanish bilinguals. Bilingualism: Language and Cognition, 2, 231-244.
Costa, A., Miozzo, M. & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? Journal of Memory and Language, 41, 365–397.
Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly-proficient bilinguals and L2 learners. Journal of Memory and Language, 50, 491–511.
Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. Brain and Language, 94, 94-103.
Costa, A., La Heij, W., & Navarrete, E. (2006b). The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition*, 9(2), 137–151.
Costa, A., Santesteban, M., & Ivanova, I. (2006). How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. Journal of Experimental Psychology: Learning, Memory and Cognition, 32 (5), 1057–1074.
Costa, A., Hernandez, M., & Sebastian-Gallés, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition, 106,* 59–86.
Costa, A., Strijkers, K., Martin, C., & Thierry, G. (2009). The time-course of word retrieval revealed by event-related brain potentials during overt speech. Proceedings of the National Academy of Sciences, 106, 21442-21446.
Costa, A., Hernández, A., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, 113, 135-149.
Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 1283–1296.
Damian, M. F., & Bowers, J. S. (2003). Locus of semantic interference in picture–word interference tasks. Psychonomic Bulletin and Review, 10, 111–117.

Damian, M. F., Vigliocco, V., & Levelt, W. J. M. (2001). Effects of semantic context in the naming of pictures and words. Cognition, 81, B77–B86.
DeBot, K., Gommans, P., & Rossing, C. (1991). L1 loss in an L2 environment: Dutch immigrants in France. In H. Seliger & R. Vago (eds.), *First language attrition*, pp. 87-98, Cambridge, UK: Cambridge University Press.
Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. Psychological Review, 93, 283–321.
Dell, G. S., Oppenheim, G. M., & Kittredge, A. K. (2008). Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. Language and Cognitive Processes, 23, 583–608.
Dell, G. S. (1990). Effects of frequency and vocabulary type on phonological speech errors. *Language & Cognitive Processes*, 5, 313-349.
Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104, 801–838.
Dhooge, E. & Hartsuiker, R.J. (2010). The Distractor Frequency Effect in Picture–Word Interference: Evidence for Response Exclusion. Journal of Experimental Psychology: Learning, Memory and Cognition, 36, 4, 878-891.
Dhooge E., Hartsuiker R. J. (2011). The distractor frequency effect in a delayed picture word interference task: further evidence for a late locus of distractor exclusion. Psychonomic Bulletin and Review, 18, 116–122.
Ehri, L. C., & Ryan, E. B. (1980). Performance of bilinguals in a picture-word interference task. Journal of Psycholinguistic Research, 9 (3), 285–302.
Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, 46, 57–84.
Finkbeiner, M., Gollan, T. H., & Caramazza, A. (2006). Lexical access in bilingual speakers: what's the (hard) problem? Bilingualism: Language and Cognition, 9, 153–166.
Finkbeiner, M., & Caramazza, A. (2006). Now you see it, now you don't: On turning semantic interference into facilitation in a Stroop-like task. Cortex, 42, 790–796.
Forster, K. I., & Forster, J. C. (2003). DMdX: A windows display program with millisecond accuracy. Behavior Research Methods Instruments and Computers, 35, 116–124.
Francis, W. S., Augustini, B. K., & Saenz, S. P. (2003). Repetition priming in picture naming and translation depends on shared processes and their difficulty: Evidence from Spanish–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1283–1297.
Glaser, W.R., & Düngelhoff, F.J. (1984). The time course of picture-word interference. Journal of Experimental Psychology: Human Perception and Performance, 10, 640-654.
Gollan, T. H., Bonanni, M. P., & Montoya, R. I. (2005). Proper names get stuck on bilingual and monolingual speakers' tip of the tongue equally often. *Neuropsychology*, 19, 278–287.
Gollan, T.H., Slattery, T.J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency Drives Lexical Access in Reading but not in Speaking: The Frequency-Lag Hypothesis. *Journal of Experimental Psychology: General*, in press.

Gollan, T. H., Weissberger, G.H., Runnqvist, E., Montoya, R., & Cera, C. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition*, 15 (03), 594-615.
Gollan, T. H., & Ferreira, V. S. (2009). Should I stay or should I switch? A cost-benefit analysis of voluntary language switching in young and aging bilinguals. Journal of Experimental Psychology: Learning, Memory & Cognition, 35, 640–665.
Gollan, T. H., Montoya, R. I., & Werner, G. (2002). Semantic and letter fluency in Spanish–English bilinguals. *Neuropsychology*, 16, 562-576.
Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., & Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, 33, 1220–1234.
Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society*, 13, 197–208.
Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, 58, 787–814.
Green, D.W. (1986). Control, activation and resource. *Brain and Language*, 27, 210–223.
Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67-81.
Griffin, Z. M., & Bock, K. (1998) Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, 38, 313-338.
Hall, M.L. (2011). Bilingual picture-word studies constrain theories of lexical selection. Frontiers in Psychology, 2, 381, doi: 10.3389/fpsyg.2011.00381.
Hartsuiker, R.J., Pickering, M. J., & Veltkamp, E. (2004). Is syntax separate or shared between languages? Cross-linguistic syntactic priming in Spanish/English bilinguals. *Psychological Science* 15, 409-414.
Hermans, D., Bongaerts, T., de Bot, K., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? Bilingualism: Language and Cognition, 1 (3), 213-230.
Hermans, D. (2004). Between-language identity effects in picture-word interference tasks: A challenge for language-nonspecific or language-specific models of lexical access? International Journal of Bilingualism, 8, 115-125.
Hermans, D. (2000). Word production in a foreign language. Ph.D. thesis, University of Nijmegen.
Hernández, M., Costa, A., Fuentes, L., Vivas, A., & Sebastian-Gallés, N. (2010).The impact of bilingualism on the executive control and orienting networks of attention. *Bilingualism: Language and Cognition*, 13, pp 315-325 doi:10.1017/S1366728909990010
Hoshino, N., & Kroll, J. F. (2008). Cognate effects in picture naming: Does cross-language activation survive a change of script? Cognition, 106, 501–511.
Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. Cognition, 100, 464-482.

Ivanova, I., & Costa A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, 127, 277–288.
Janssen, N., Schirm, W., Mahon, B.Z., & Caramazza, A. (2008). The semantic interference effect in the picture-word interference paradigm: Evidence for the response selection hypothesis. Journal of Experimental Psychology: Learning, Memory, & Cognition, 34, 249-256.
Köpke, B. (2007). Language attrition at the crossroads of brain, mind, and society. In B. Köpke, M. S. Schmid, M. Keijzer & S. Dostert (eds.), *Language attrition: Theoretical perspectives*. Amsterdam: John Benjamins, 9–37.
Köpke, B. & Schmid, M. S. (2004). Language attrition: The next phase. In B. Schmid, M. S. Köpke, M. Keijzer & L. Weilemar (eds.), *First language attrition: Interdisciplinary perspective on methodological issues*, pp. 1–43. Amsterdam & Philadelphia: John Benjamins.
Kroll, J.F., Bobb, S.C., & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. Bilingualism: Language and Cognition, 9, 119–135.
Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33, 149-174.
Kroll, J. F., Bobb, S. C., Misra, M. & Guo, T. (2008). Language selection in bilingual speech: Evidence for inhibitory processes. *Acta Psychologica*, 128, 416-430.
La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J. F. Kroll & A. M. B. De Groot (eds.), Handbook of bilingualism: Psycholinguistic approaches, pp. 289–307. New York: Oxford University Press.
La Heij, W., Starreveld, P. A., & Steehouwer, L. C. (1993). Semantic interference and orthographic facilitation in definition naming. Journal of Experimental Psychology: Learning, Memory, and Cognition, 19, 352–368.
La Heij, W., Dirkx, J., & Kramer, P. (1990). Categorical interference and associative priming in picture naming. British Journal of Experimental Psychology, 81, 51-525.
Lee, M.W., & Williams, J.N. (2001). Lexical access in spoken word production by bilinguals: Evidence from the semantic competitor priming paradigm. *Bilingualism: Language and Cognition*, 4, 233–248.
Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioural and Brain Sciences*, 22, 1–75.
Levy, B. J., & Anderson, M. C. (2002). Inhibitory processes and the control of memory retrieval. *Trends in Cognitive Science*, 6, 299–305.
Levy, B. J., Mc Veigh, N., Marful, A., Anderson, M. C. (2007). Inhibiting your native language. The Role of Retrieval-Induced Forgetting During Second-Language Acquisition. *Psychological Science*, 18 (1), 29-34.
Li, H. & MacWhinney, B. (2011). Semantic competitor priming within and across languages: The interplay of vocabulary knowledge, learning experience and working memory capacity. Bilingualism: Language and Cognition, 14, 433-443.
Loebell, H., & Bock, K. (2003). Structural priming across languages. *Linguistics*, 41 (5), 791–824.
Luo, L., Luk, G. & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition*, 114, 29–41.
MacLeod, M.D & Macrae, C.N. (2001) Gone but not forgotten: the transient nature of retrieval-induced forgetting. *Psychological Science*, 12, 148–152.

Macrae, C.N. & MacLeod, M.D. (1999). On recollections lost: when practice makes imperfect. *Journal of Personality and Social Psychology*, 77, 463–473
MacWhinney, B. (2005). A unified model of language acquisition. In J. Kroll and A. DeGroot (eds) *Handbook of Bilingualism*, Oxford: Oxford University Press, 49-67.
Mägiste, E. (1984). Stroop tasks and dichotic translation: The development of interference patterns in bilinguals. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 304-315.
Mägiste, E. (1985). Development of intra- and inter-lingual interference in bilinguals. Journal of Psycholinguistic Research, 14, 137–154.
Mahon, B.Z., Costa, A., Peterson, R., Vargas, K.A., Caramazza, A. (2007). Lexical Selection is Not by Competition: A Reinterpretation of Semantic Interference and facilitation Effects in the Picture-Word Interference Paradigm. *Journal of Experimental Psychology: Learning Memory and Cognition*, 33, 503-535
Martin, R.C., Crowther, J.E., Knight, M., Tamborello, F.P., & Yang, C.L. (2010). Planning in sentence production: evidence for the phrase as a default planning scope. *Cognition*, 116, 77-92.
Meuter, R.F.I. & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. Journal of Memory and Language, 40, 25-40.
Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distracter frequency in the picture–word interference paradigm. Journal of Experimental Psychology: General, 132, 228–252.
Navarrete, E., Mahon, B.Z. & Caramazza, A. (2010). The cumulative semantic cost does not reflect lexical selection by competition. Acta Psychologica, 134 (3), 279-289.
Oppenheim, G.M., Dell, G.S., Schwartz, M.F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. Cognition, 114, 227–252.
Paradis, M. (2007). L1 attrition features predicted by a neurolinguistic theory of bilingualism. In B. Köpke, M. S. Schmid, M. Keijzer & S. Dostert (eds.), *Language Attrition: Theoretical perspectives*, pp. 121 – 134. John Benjamins Publishing Company, Amsterdam.
Philipp, A.M. & Koch, I. (2009). Inhibition in language switching: What is inhibited when switching among languages in naming tasks? Journal of Experimental Psychology: Learning, Memory, & Cognition, 35, 1187-1195.
Quene H, van den Bergh H. Examples of mixed-effects modeling with crossed random effects and with binomial data. Journal of Memory and Language, 59, 413–25.
Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. Cognition, 42, 107–142.
Rosinksi, R. R., Golinkoff, R. M., & Kukish, K. S. (1975). Automatic semantic processing in a picture-word interference task. Child Development, 46, 247-253.
Runnqvist, E. & Costa, A. (in press). Is retrieval-induced forgetting behind the bilingual disadvantage in speech production? Bilingualism: Language and Cognition, DOI: 10.1017/S1366728911000034.
Sadat, J., Martin, C.D., Alario, F.X., & Costa, A. (2011). Characterizing the bilingual disadvantage in noun phrase production. *Journal of Psycholinguistic Research*, DOI: 10.1007/s10936-011-9183-1

Sandoval, T.C., Gollan, T.H., Ferreira, V.S. & Salmon, D.P. (2010).What causes the bilingual disadvantage in verbal fluency? The dual-task analogy. *Bilingualism: Language and Cognition*, in press.
Schmid, M. S. (2002). *First language attrition, use and maintenance: The case of German jews in anglophone countries.* Benjamins: Amsterdam.
Schmid, M. S. (2010). Languages at play: The relevance of L1 attrition to the study of bilingualism. *Bilingualism: Language and Cognition*, 13, Special Issue 01, 1-7.
Schriefers, H., Meyer, A.S., & Levelt, W.J.M. (1990). Exploring the time course of lexical access in language production: Picture-word interference studies. Journal of Memory and Language, 29, 86-102.
Seliger, H.W., & Vago, R.M. (1991). The study of first language attrition: an overview. In H. W. Seliger & R. M. Vago (eds.), *First language attrition*, pp. 3-16. Cambridge, UK: Cambridge University Press.
Shin, J.A., & Christianson, K. (2009). Syntactic processing in Korean-English bilingual production: Evidence from cross-linguistic structural priming. Cognition, 112(1), 175–180.
Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 686–698.
Starreveld, P. A., & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture-naming. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 896–918.
Strijkers, K., Costa, A., & Thierry, G. (2010). Tracking lexical access in speech production: Electrophysiological correlates of word frequency and cognate effects. Cerebral Cortex, 20 (4), 912-928.
Szmrecsanyi, B. (2009). The great regression: genitive variability in Late Modern English news texts. In K. Börjars, D. Denison, and A. Scott (Eds.), Morphosyntactic categories and the expression of possession. Amsterdam, Philadelphia: Benjamins.
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.
Verhoef, K. M. W., Roelofs, A., & Chwilla, D. J. (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. Cognition, 110, 84–99.
Vitkovitch, M., & Humphreys, G. W. (1991). Perseverant responding in speeded naming of pictures: It's in the links. Journal of Experimental Psychology: Learning, Memory, and Cognition, 17, 664-680.
Wheeldon, L. R. & Monsell, S. (1994). Inhibition of spoken word production by priming a semantic competitor. Journal of Memory and Language, 33, 332–356.
Xiao, R., McEnery, T. & Qian, Y. (2006). Passive constructions in English and Chinese: A corpus-based contrastive study. *Languages in Contrast*, 6(1) 109-149.
Zeelenberg, R. & Pecher, D. (2003). Evidence for long-term crosslanguage repetition priming in conceptual implicit memory tasks. *Journal of Memory and Language*, 49, 80–94.

INDEX: