Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners

Albert Costa\textsuperscript{a,}\textsuperscript{*} and Mikel Santesteban\textsuperscript{b}

\textsuperscript{a} Departament de Psicologia Bàsica, Universitat de Barcelona, P. de la Vall d’Hebron, 171 08035, Barcelona, Spain
\textsuperscript{b} EHU, University of the Basque Country, Spain

Received 1 December 2003; revision received 11 February 2004

Abstract

Five experiments are reported in which the picture naming performance of bilingual speakers in a language-switching task was explored. In Experiment 1, Spanish learners of Catalan and Korean learners of Spanish were asked to perform a switching task between their first and dominant language (L1, Spanish or Korean) and their second language (L2, Catalan or Spanish). For these two groups switching from the weaker language (L2) to the more dominant language (L1) was harder than vice versa. This asymmetrical switching cost was not present when highly proficient Spanish–Catalan bilinguals performed the task either in their two dominant languages (Experiments 2 and 3) or in their dominant language (L1) and in their much weaker language (L3 English; Experiment 4). Furthermore, highly proficient bilinguals showed faster naming latencies in their weaker languages (L2 and L3) than in their dominant language (L1). Experiment 5 tested whether a bias in the triggering of lexicalization is at the basis of such a difference. Together these results reveal that the switching performance of highly proficient bilinguals does not seem to be subject to the same mechanisms as that of L2 learners.

© 2004 Elsevier Inc. All rights reserved.

Introduction

One of the most remarkable abilities of bilingual speakers is that of separating their two languages during the production of speech. Although the speech of highly proficient bilinguals in their second language (L2) often carries traces of the first language (L1) (e.g., L1 accent, L1 syntactic structures; Flege, MacKay, & Meador, 1999; Pallier, Colomé, & Sebastián-Gallés, 2001; Yeni-Komshian, Flege, & Liu, 2000), it rarely exhibits L1 lexical intrusions (Poulisse, 1999; Poulisse & Bongaerts, 1994). That is, when necessary, these bilinguals are quite competent at selecting and producing words from only one of their lexicons, both in L1 and, more remarkably, in their L2. What are the mechanisms controlling lexical access in bilingual speech production? A clear understanding of how two languages are handled by one mind requires, among other things, that we investigate the mechanisms that allow bilingual speakers to restrict their lexicalization process to only one language. In this article, we address this issue by investigating the naming performance of several groups of bilingual speakers in a language switching task. The main objective of this research is to assess how L2 proficiency level affects the processes of lexical selection in speech production. We will do so by comparing the patterns of language switching performance of various groups of bilinguals that vary in their proficiency levels.

The Introduction is structured as follows. First, we discuss the general functional architecture and dynamics
of the speech production system. Second, we introduce the hypothesis that lexical access in bilingual speakers involves inhibitory control, and discuss how the language switching task can provide useful information to test such a hypothesis. Before presenting the overview of the experiments, we put forward the hypothesis that the language switching performance may vary depending on the bilingual’s proficiency levels.

**Lexical access in bilingual speech production**

A central stage in language production is that in which speakers retrieve the words from the lexicon that match their communicative intention. The process by which this is achieved is often referred to as lexical selection (Caramazza, 1997; Dell, 1986; Levelt, 1989, 2001; Levelt, Roelofs, & Meyer, 1999). A selection mechanism is needed because several lexical representations are activated due to spreading activation from the semantic system to the lexical level. Thus, any representation activated at the conceptual level spreads a proportion of its activation to its corresponding lexical node. In this scenario, the semantic system activates not only the word that matches the intended meaning but also other semantically related items. That is, when naming the picture of a dog, not only the lexical node “dog” is activated, but also other related words, such as “cat” and “bark.” The lexical selection mechanism is in charge of deciding which of the activated lexical items needs to be prioritized for further processing. It is widely accepted that the level of activation of lexical nodes is the critical variable for deciding which element is to be selected. Thus, in general, the lexical selection mechanism would pick out the word with the highest level of activation which, in normal error-free production, corresponds to the intended meaning. However, some models of lexical access assume that this mechanism is also sensitive to the level of activation of non-target—but activated—lexical nodes that act as competitors. That is, the ease with which a word is selected depends on its level of activation relative to that of other lexical items (e.g., Roelofs, 1992).

When implementing the lexical selection mechanism in the context of bilingual speech production, there are two questions that become relevant. First, does the semantic system activate the two lexicons of a bilingual? Second, do the lexical nodes of the non-response language (the language not intended for production) act as competitors? The first question has a positive answer, and present models of lexical access assume that during the course of lexicalization in one language (e.g., L2), the lexical nodes of both languages of a bilingual receive activation from the semantic system (e.g., Colomé, 2001; Costa, Caramazza, & Sebastián-Gallés, 2000; Costa, Colomé, Gomez, & Sebastián-Gallés, in press; De Bot, 1992; Gollan & Kroll, 2001; Hermans, Bongaerts, de Bot, & Schreuder, 1998; Poulisse, 1999).

More controversial is the answer to the second question. Some models of lexical access assume that the lexical selection mechanism is language-specific (Costa, in press; Costa & Caramazza, 1999; Costa, Miozzo, & Caramazza, 1999; Roelofs, 1998), in the sense that it only considers the activation-levels of words in the intended language. According to this idea, lexical intrusions from the non-intended language would be prevented since those words would not be included in the pool of possible candidates for production, and therefore they will not be able to interfere during lexical access. The notion that lexical selection can be sensitive to specific properties of lexical nodes, and can use them to guide selection, has already been postulated in models of monolingual lexical access. For example, in Dell’s model (Dell, 1986), the lexical selection mechanism is sensitive to the grammatical class of lexical items. That is, if the speaker wants to produce a noun the selection mechanism would consider for selection only lexical items corresponding to nouns (see Mahon, Costa, Shapiro, & Caramazza, 2002; Pechmann & Zerbst, 2002, for evidence consistent with such an assumption).

In contrast, other models of bilingual lexical access assume that the lexical selection mechanism is insensitive to the language in which the speaker intends to express her ideas. In such a framework, the speaker would consider for selection all activated lexical nodes, irrespective of the language to which they belong, and successful selection of the proper lexical node (i.e., in the correct language) is achieved by creating a differential level of activation in the two lexicons of a bilingual. Thus, the question here is: How does the system produce an imbalance of activation between the two lexicons? One way to do so is to assume that the semantic system activates words in the intended language to a larger extent than words in the non-response language (La Heij, in press; Poulisse & Bongaerts, 1994). The second solution is to postulate that lexical access in bilingual speakers entails the reactive inhibition of lexical items belonging to the non-response language (Green, 1986, 1998; Hermans et al., 1998).

**Inhibitory control in bilingual speech production and language switching**

In the present study, we test the second of these implementations. We address whether inhibitory control, per se, can account for lexical access in highly proficient bilingual speakers. We do so by studying the properties of one experimental effect that has often been interpreted as supporting the notion that lexical access entails inhibitory processes; the asymmetrical switching cost reported by Meuter and Allport (1999) (see below).

The most articulated model of inhibitory control in bilingual speech production is that proposed by Green (1998). Among several assumptions made by this model
L1 was more costly than vice versa (see also Janssen, 1999; switching costs were asymmetrical: to switch from L2 to L1 incurred a time cost. However, the magnitude of the inhibition applied to one language depends on the speaker’s proficiency level in that language. In other words, when speaking in L1, not much inhibition is required for the less dominant language (L2) because it is assumed that the baseline level of activation of L2 lexical items is lower than that of L1 lexical items. However, when speaking in the less dominant language (L2), L1 representations must be strongly inhibited in order to ensure that L2 lexical items are selected. The second assumption crucial for present purposes refers to the time required for inhibition to be overcome. It is assumed that the suppression (or inhibition) of the activation of a given language may exert an influence on subsequent production events in which words from the suppressed lexicon need to be retrieved (see Allport, Styles, & Hsieh, 1994). Accordingly, retrieving words from a lexicon that has just been inhibited would be relatively difficult, since it would take relatively more time for that inhibition to be overcome. Therefore, the more inhibition applied to a given lexicon, the harder it will be to overcome such suppression on a subsequent trial (see Tipper, 2001; for discussion of negative priming). This sort of model can be considered language non-specific, given that the computation responsible for lexical selection considers the activation levels of all lexical nodes, irrespective of the language to which they belong.

There is one set of results that has been argued to support the two assumptions made by the IC model. Meuter and Allport (1999) conducted a language switching experiment in which bilingual speakers were asked to name aloud series of lists containing Arabic digits (from 1 to 9) either in L1 or in L2. The language in which a given number had to be named was signaled by the background color of the screen (i.e., if blue name the digit in L1, if red name the digit in L2). Experimental trials were divided into switch and non-switch trials. A non-switch trial required a response in the same language as the immediately preceding trial, while a switch trial required a response in a different language. Given that responses were produced in both L1 and L2, there were four different types of trials: Switch to L1, Switch to L2, Non-Switch in L1, Non-Switch in L2. As expected, naming latencies in switch trials were slower than in non-switch trials, revealing that switching between tasks (name in language X, name in language Y) incurred a time cost. However, the magnitude of the switching cost was larger for L1 than for L2. That is, the switching costs were asymmetrical: to switch from L2 to L1 was more costly than vice versa (see also Janssen, 1999). The asymmetrical switching cost, which at first glance seems paradoxical, finds a ready explanation in the framework of the IC model and in the Task Set Inertia hypothesis (e.g., Allport et al., 1994; Allport & Wylie, 1999). As Meuter and Allport argued, the magnitude of inhibition applied to L1 when speaking in L2 must be greater than vice versa in order to allow successful selection of L2 lexical items (assumption 1 of the IC model). That is, the relative strength of the two tasks at hand (namely in L1 or in L2) would have an effect on the strength with which each of the two languages is inhibited. The authors further assumed that the inhibition exerted in one language has effects on the subsequent trial. As a consequence, to retrieve L1 words in a switch trial will be relatively hard because the system has to overcome the large inhibition that was applied to that lexicon in the immediately preceding trial (assumption 2 of the IC model). However, to retrieve L2 words in a switch trial will be relatively less difficult since the L2 lexicon was not as strongly suppressed in the preceding trial.

Evidence converging on the assumption that access to L2 representations involves the active suppression (or inhibition) of L1 was reported by Jackson, Swainson, Cunnington, and Jackson (2001) who recorded participants’ event-related electrical potentials (ERPs) in a language switching task. Switch trials in L2 were associated with an enhanced negative deflection in the ERP at frontal sites compared to L2 non-switch trials. However, such a difference was not observed for L1 trials. Given that this component seems to index the presence of response suppression (e.g., Jackson, Jackson, & Roberts, 1999), the authors interpreted its asymmetrical presence in L2 as revealing the larger suppression of L1 in the switch trials (see also Hernandez, Martinez, & Kohnert, 2000; Hernandez, Dapretto, Mazziotta, & Bookheimer, 2001; Price, Green, & Von Studnitz, 1999; for brain imaging and electrophysiological studies in which language switching has been interpreted as revealing inhibitory processing). Further support for the notion that the asymmetrical switching cost reveals different degrees of language dominance comes from the study of Hernandez and Kohnert (1999, personal communication). In that study, a larger switching cost was observed in the dominant language (English), despite the fact that for those bilinguals (Spanish–English) Spanish was the earlier acquired language.

Note, however, that the above explanation of the existence of asymmetric language switching costs is a bit more complex. There are two components contributing to the existence of asymmetrical switching costs. The first is the time it takes for the speaker to implement the new task goal, or in Green’s model, to select the right “language schema” (see also Allport & Wylie, 2000; Meiran, 2000; Ruthruff, Remington, & Johnston, 2001;
Sohn & Anderson, 2001 for similar arguments in general task-switching models). This component is common for both directions of switching (from L1 to L2 and vice versa), and reflects the intrinsic time it takes for speakers to reset their experimental goal in switch trials (from naming in language X to naming in language Y). This switching cost is common to many other switching tasks, and is supposed to reflect processes outside the lexical system (e.g., Von Studnitz & Green, 2002). The second component refers to the inhibition applied to the lexical representations of the language that is not involved in a given naming trial. This component is assumed to be responsible for observing the asymmetrical switching costs in the language switching task. This is because in the switching task L1 lexical representations are more inhibited than L2 lexical representations. Thus, while the first component of the switching cost accounts for the presence of switching costs per se, the second component accounts for the presence of the asymmetrical cost.

Further post hoc analyses conducted by Meuter and Allport (1999) supported the notion that the asymmetrical switching costs stem from differences in the L2–L1 proficiency levels. In particular, the asymmetrical switching costs were somewhat dependent on the bilinguals’ L2 proficiency: the higher the L2 proficiency, the smaller the asymmetry. Although, such differences among groups did not turn out to be significant, the trend was in the direction predicted by the IC model, and led the authors to speculate that the asymmetrical switching cost may disappear for highly proficient bilinguals. This is because for these bilinguals, the difference in proficiency levels between the two languages would be minimal, and therefore the amount of inhibition applied to both languages would be similar (see Monsell, Yeung, & Azuma, 2000; for a discussion of the conditions in which asymmetrical switching costs are observed).

Language switching and bilingual proficiency levels

In the present article, we adopt the explanation given by Meuter and Allport (1999) of asymmetrical switching costs and test whether such an explanation can account for the switching performance of highly proficient early bilinguals.

To contrast the performance of high and low proficient bilinguals is important, because L2 proficiency seems to be one of the most relevant factors for predicting bilingual speech performance. For example, proficiency correlates negatively with the number of lexical intrusions in speech production (Poulisse & Bongaerts, 1994). That is, highly proficient bilinguals suffer fewer L1 intrusions than low-proficient bilinguals when speaking in their L2. Also, proficiency seems to be one of the factors governing the representation of L2 in the brain (e.g., Perani et al., 1998). Thus, in the experiments to be reported we pay special attention to highly proficient bilinguals that have acquired their two languages early in life. The issue of whether L2 proficiency or L2 age of acquisition is at the basis of the performance of these speakers will not be directly addressed in our study, and we defer discussion of the relevance of these two variables until the General Discussion.

The differences between various types of bilingual speakers may be explained in terms of a difference in quantitative or qualitative terms. The difference may be revealing the different degrees of mastery with which the same lexical access mechanism is handled by different bilingual populations. For example, the fact that proficiency correlates negatively with language intrusions, may be accounted for by assuming that proficiency correlates with better inhibitory control. Also, such an effect could be revealing a better control of the amount of activation that the semantic system sends to the two lexicons. In other words, becoming a highly proficient bilingual would be a consequence of applying the same mechanisms available to less-proficient bilinguals in a more skilled way. However, a difference between high and low-proficient bilinguals could also be revealing a qualitative difference between the two populations. Thus, one may speculate on whether lexical selection in bilingual speakers entails different types of cognitive processes depending on their proficiency level. In other words, a difference in the performance of highly proficient bilinguals and less fluent bilinguals may stem not from “better” processing but from “different” processing.

Note that the aim of our research is not to investigate issues related to task switching per se, but rather to assess how experiments of this sort can inform us about the way in which bilingual speakers produce fluent speech (see Monsell, 2003, for a recent review of the task switching paradigm; and Meuter, in press, for a review on language switching). Nevertheless, this research can also inform us about the validity of the switching paradigm to study issues related to language processing in bilingual speakers, if it can be shown that the paradigm is sensitive to the linguistic particularities of the speakers performing the task.

Overview of the experiments

The experiments presented in this article have two main goals: first, to replicate the observations of Meuter and Allport (1999); second, and more importantly, to test whether L2 proficiency affects the pattern of switching performance. Experiment 1 is devoted to the first objective while Experiments 2–5 to the second.

In Experiment 1, we investigate the switching performance of two different bilingual populations (Spanish–Catalan and Korean–Spanish). The speakers in this experiment could be considered to be L2 learners.
In Experiment 2, we analyze the performance of highly proficient Spanish–Catalan bilinguals when switching between their two dominant languages (Spanish and Catalan). In Experiment 3, we replicate the observations made in Experiment 2 under somewhat different experimental conditions (e.g., larger set of experimental materials). Experiment 4 explores the switching performance of highly proficient Spanish–Catalan bilinguals when the task involves switching between their dominant language L1 (Spanish) and their much weaker language L3 (English). Finally, Experiment 5 extends the observations made in Experiments 2 and 3 to another experimental situation that allows participants more time to prepare for the language switch.

Despite the potential limitations of the specific properties of the design used by Meuter and Allport (1999) for drawing conclusions about the performance of bilingual speakers in other contexts (e.g., the large number of repetitions of items), and to be able to compare both studies, we keep our experimental design, in the relevant dimensions (e.g., % of switching trials, % of responses in L1 and L2, etc.) as similar as possible to theirs. There are, however, several important modifications. First, the populations tested in each of our experiments are homogeneous (regarding the languages of the speakers), while the population tested in Meuter and Allport’s experiment was not (French–English, German–English, etc.). Second, the number of trials, and the number of repetitions of the experimental stimuli, is much smaller in our study (from 2000 trials to 950). Third, to avoid the use of phonologically similar translations (cognates) we used pictures of common objects rather than Arabic numerals as experimental stimuli.

**Experiment 1: Replicating the asymmetrical switching costs in L2 learners**

This experiment aims at replicating the asymmetrical switching costs reported by Meuter and Allport (1999) for low-proficient bilinguals. We do so by investigating the performance of two groups of speakers who have been exposed to their L2 relatively late in life and who can be considered L2 learners. The languages of the speakers in the two groups were different. In Group 1, we tested Spanish–Catalan participants and in Group 2, Korean–Spanish participants. Given the results of Meuter and Allport (1999), we expect the switching cost in L1 to be larger than in L2, for both groups of participants.

**Method**

**Participants**

Twenty-four participants took part in the experiment. Participants in Group 1 were native speakers of Spanish that were learning Catalan for an average of 1.5 years. These participants were receiving formal training in Catalan as a second language. Participants in Group 2 were native speakers of Korean that were learning Spanish for an average of 4 years. All participants were living in Barcelona at the time of testing, and were therefore, regularly exposed to Catalan and Spanish. Participants in both groups were considered to be L2 learners (see language proficiency self-assessment and history in Appendix A).

**Materials**

The materials used for both groups of participants were similar. For Group 1, 10 pictures of common objects with non-cognate names were selected (see Appendix B for a description). For Group 2, we used 8 out of the 10 pictures plus two new ones. All pictures' names were non-cognates.

Participants were instructed to choose the language of the response according to the color of the picture (red or blue). The assignment of color cue to response language was counterbalanced across participants. Half of the participants were instructed that “red” indicated “respond in Spanish (or Korean for Group 2)” and “blue” indicated “respond in Catalan (or Spanish for Group 2),” and the other half received the reverse assignments. Pictures were presented in short sequences (“lists”) ranging in length unpredictably from 5 to 14 trials. There were two types of trials: (a) trials in which the language of the response (either L1 or L2) was the same as the trial immediately before (non-switch trials); (b) trials in which the language of the response (either L1 or L2) was different than that used in the preceding trial (switch trials).

The lists varied in the number of switch trials (from 0 to 4 switching trials). A total of 400 such lists were constructed. These lists varied on the specific sequence: (a) L1 and L2 responses; and (b) switch and non-switch trials. Each participant was presented with 100 of such lists that varied in length (from 5 to 14 trials). In total, each participant was presented with 950 trials (70% non-switch and 30% switch). In half of the non-switch trials, participants were asked to name the picture in L1 (333 trials) and in the other half in L2 (333 trials). The same applied for the switch trials. Therefore, participants used their L1 and L2 the same number of times during the course of the experiment (475 responses in each language). For each subject, 50 of the 100 lists started with a picture to be named in L1 and 50 with a picture to be named in L2.

Each picture was presented 95 times during the experiment. The assignment of each specific picture to each trial was left random, but varied for each participant. That is, a given picture could be named more often in L1 than in L2 and could appear more often in a switch trial than in a non-switch trial for one participant. However,
for another participant this distribution changed, given that the assignment of the pictures to the different trial types varied randomly across participants. In lists of 10 trials or fewer no picture could appear twice, and in lists from 11 to 14 items the same picture could appear no more than twice. There were at least two different items between the first and second appearance of the same picture.

Procedure

Participants were tested individually in a soundproof room. They were asked to name the pictures as fast as possible while trying not to make errors. They were informed that the language in which a given picture had to be named was determined by the color in which the picture appeared, and that in a given list of trials, pictures with different colors could be presented. Before the experiment proper, participants were familiarized with the names of the pictures in the two languages. A list of trials had the following structure: (1) a blue or red circle along with the word CATALÀ (Catalan) or ESPANYOL (Spanish) written below was presented for 2000 ms (the words were Spanish or Korean for Group 2). This circle (and the word) indicated the language in which the first picture of the list had to be named, and therefore we can consider the first trial of a list as a non-switch trial; (2) the first picture of a list appeared and remained on the screen for 2000 ms or until participants responded; (3) a blank interval of 1150 ms; (4) the next picture was presented, and the cycle was repeated until the end of the list; (5) after the presentation of the last picture of the list an asterisk was presented for 1000 ms signalling the end of the list. The experiment started with the presentation of 6 training lists.

Data analysis

Two main variables were considered in the statistical analyses: “Response Language” (L1 or L2), and “Type of Trial” (Switch vs Non-switch). Error rates and naming latencies were submitted to two analyses of variance (ANOVAs). The analyses were conducted for each Group of participants separately. Although, analyses by items are not usually reported in this type of experiment (Meuter & Allport, 1999), for the sake of completeness we present them here. However, given the small number of items, and that we assigned them randomly to each condition, caution has to be exercised when interpreting the item analyses.

Two types of responses were scored as errors: (a) verbal disfluencies (stuttering, utterance repairs, production of nonverbal sounds that triggered the voice key); (b) recording failures. Also, naming latencies exceeding 3 SD from a given participant’s mean and trials in which participants produced a different name from that designated by the experimenter were discarded from the analyses.

Results

Group 1: Spanish speaking learners of Catalan

Following the criteria presented above, 5.8% of the trials (errors: 4.5%; outliers: 1.3%) were excluded from the analyses. No significant effects were observed in the error analyses.

In the analysis of naming latencies, the main effects of “Response Language” \(F(1,11) = 7.04; MSE = 2143.20; \ p = .022; F(2,1,9) = 5.9; MSE = 1928.12; \ p = .038\) and “Type of Trial” \(F(1,11) = 53.80; MSE = 631.47; \ p = .001; F(2,1,9) = 221.45; MSE = 130.45; \ p = .001\) were significant. More importantly, the interaction between the two variables was also significant \(F(1,11) = 8.83; MSE = 437.97; \ p = .014; F(2,1,9) = 15.39; MSE = 213.94; \ p = .003\), revealing that the magnitude of the switching cost was different for the two languages (L1: 71 ms; L2: 35 ms). Nevertheless, the switching cost was significant for both languages (two tailed \(t\) tests; L1: \(t(11) = 6.38; \ p = .001\); L2 \(t(11) = 4.78; \ p = .001\) (see Fig. 1).

![Figure 1](image-url)
Group 2: Korean speaking learners of Spanish

We discarded 6.8% (errors: 5.2%; outliers: 1.6%) of the data points. No significant effects were observed in the error analyses.

In the analysis of naming latencies, the only significant main effect was that of the variable “Type of Trial” ($F(1, 11) = 38.40; MSE = 1184.99; p = .001; F2(1, 9) = 51.01; MSE = 590.82; p = .001). Importantly, the interaction between the variables “Type of Trial” and “Response Language” was significant ($F(1, 11) = 17.46; MSE = 261.27; p = .002; F2(1, 9) = 9.76; MSE = 288.9; p = .012), revealing that the magnitude of the switching cost was different for the two languages (L1: 81 ms; L2: 42 ms). Nevertheless, the switching cost was significant for both languages (two tailed t tests; L1: $t(11) = 7.01; p = .001; L2 t(11) = 4.07; p = .002) (see Fig. 1B).

Post hoc analyses combining the results of both groups of participants revealed that the crucial interaction reflecting asymmetrical switching costs (“Type of Trial” × “Response Language”) did not interact with the factor “Group of Participants” (both $F_s < 1$), indicating that the magnitude of the asymmetrical switching cost was not different in the two groups.

Discussion

The results of Experiment 1 were clear: the magnitude of the switching cost was larger for L1 than for L2. That is, there was an asymmetrical switching cost depending on the response language, replicating Meuter and Allport’s (1999) observation. Although, this asymmetrical switching cost was observed for both groups of participants, the shape of the interaction was somewhat different. While in Group 1 responses for non-switch trials were somewhat faster (but non-significantly faster) for L2 than for L1, the opposite was true for Group 2. Interestingly, both types of interaction have already been observed in similar studies. For example, while the shape of the interaction for Group 1 resembles that observed by Janssen (1999), the performance of participants in Group 2 resembles that reported by Meuter and Allport (1999). A third type of interaction has been found by Jackson and collaborators (2001) in which naming latencies were faster for L1 than for L2 in the non-switch trials and no differences among languages were found in the switch trials. At present, the source of the variation across studies is unclear (see Jackson et al., 2001; Yeung & Monsell, 2003; for a discussion of this issue). What is important in the present context, however, is that in all these studies the magnitude of the switching cost was larger for L1 than for L2.

Having established the reliability of the asymmetrical switching cost and its reproducibility with another naming task (picture naming rather than Arabic digit naming), we proceed to assess whether the presence of such an asymmetry is related to the different degree with which the two languages of a bilingual are inhibited. Specifically, does the extent to which a language is inhibited depend on the L2 proficiency level? Recall, that according to Meuter and Allport (1999) the difference in the magnitude of the switching cost is due to the larger inhibition applied to L1 in comparison to that applied to L2. This difference in the amount of inhibition reflects the difference in the proficiency levels between L1 and L2. There is a clear prediction that one can derive from this set of assumptions: when the difference between L1 and L2 proficiency levels is small (as in the case of highly proficient bilinguals), then a similar degree of inhibition should be applied to the two languages and therefore the magnitude of the switching cost should be similar in both directions.

Experiment 2: Language switching costs in highly proficient bilinguals

The objective of this experiment is to assess whether a reduction in the difference between L1 and L2 proficiency levels leads to a reduction in the magnitude of the asymmetrical switching cost. If the interpretation of the asymmetrical switching cost advanced by Meuter and Allport (1999) is correct, then we should expect a reduction (or even an elimination) of such an asymmetry when highly proficient Spanish–Catalan bilinguals perform the task in their L1 and L2. That is, to switch from L2 to L1 should be as difficult as to switch from L1 to L2.

Method

Twelve native speakers of Spanish who were highly proficient speakers of Catalan took part in the experiment (see Appendix A). The same materials and procedure used in Experiment 1 (for Group 1) were employed here.

Results and discussion

Following the same criteria as in Experiment 1, 4.9% (errors: 3.1%; outliers: 1.8%) of the data points were excluded from the analyses.

The only significant effect in the error analysis was that of the “Type of Trial” ($F(1, 11) = 8.97; MSE = 2.18; p = .012; F2(1, 9) = 9.07; MSE = 1.8; p = .015). Switch trials (5.7%) led to more errors than non-switch trials (4.5%).

In the analysis of naming latencies, the main effects of the “Response Language” ($F(1, 11) = 6.95; MSE = 650.40; p = .023; F2(1, 9) = 4.47; MSE = 854.35; p = .064) and “Type of Trial” ($F(1, 11) = 134.27; MSE = 199.52; p = .001; F2(1, 9) = 132.37; MSE = 173.69;
The results of this experiment showed that highly proficient bilinguals suffer the same switching cost in L1 and L2. These results contrast sharply with those reported in Experiment 1 for L2 learners, for whom switching into L1 was harder than switching into L2. Given that we use the same materials and languages in Experiments 1 (Group 1) and 2, we conducted a joint analysis to determine whether this difference between groups was statistically significant.

In this analysis the main effects of the variables: “Response Language” \( F(1, 22) = 12.93; \text{MSE} = 1396.80; p = .002 \); \( F(2, 1, 9) = 5.53; \text{MSE} = 2567.83; p = .043 \); “Type of Trial” \( F(1, 22) = 145.73; \text{MSE} = 415.49; p = .001 \); \( F(2, 1, 9) = 231.89; \text{MSE} = 223; p = .001 \); and “Group of Participants” \( F(1, 22) = 4.46; \text{MSE} = 34190.90; p = .046 \); \( F(2, 1, 9) = 740.13; \text{MSE} = 177.95; p = .001 \) were significant. Only one two-way interaction was significant: “Response Language” × “Type of Trial” \( F(1, 22) = 6.35; \text{MSE} = 304.32; p = .019 \); \( F(2, 1, 9) = 11.15; \text{MSE} = 129.5; p = .009 \). Crucial for our purposes here, the three-way interaction was significant (“Response Language” × “Type of Trial” × “Group of Participants”); \( F(1, 2, 2) = 6.35; \text{MSE} = 304.32; p = .019 \); \( F(2, 1, 9) = 9.91; \text{MSE} = 187.89; p = .012 \). This interaction reflects the existence of an asymmetrical switching cost for one group of participants (L2 learners [Group 1 in Experiment 1]) but not for the other group (highly proficient bilinguals in Experiment 2).

This pattern of results is consistent with the notion that the asymmetrical switching cost is related to the different degree with which the two languages of a bilingual are inhibited, and that the extent to which a language is inhibited depends on the L2 proficiency level. Thus, when the difference in proficiency between the two languages of a bilingual is large (L2 learners), more inhibition is applied to L1 than to L2, resulting in the asymmetrical switching cost. When the difference in proficiency is small (highly proficient bilinguals), a similar degree of inhibition is applied to the two languages and symmetrical switching costs are observed.

However, there is one result in this Experiment that came about as a surprise: highly proficient bilinguals named the pictures in their L1 more slowly than in their L2. If the amount of inhibition applied to L1 and L2 is similar, as suggested by the symmetrical switching costs, one should expect similar naming latencies in the two languages or, if anything, faster naming latencies in the more dominant L1. We defer further discussion of this issue to the Discussion section of Experiment 4.

One could argue that the lack of an asymmetrical switching cost in this experiment is not very informative since it stems from a null result (e.g., the lack of an interaction between “Type of Trial” and “Response Language”). However, it is important to keep in mind that this lack of an interaction is observed in the context of two main effects (“Type of Trial” and “Response Language”) and in the context of other experiments in which such an interaction is present (Experiment 1 for Groups 1 and 2). Therefore, it is unlikely that the absence of an interaction in the present experiment is due to insufficient power in our experimental design, and that we are actually committing a type II statistical error. Nevertheless, given the relevance of the observed results for models of language production in bilingual speakers, in Experiment 3 we replicate the pattern of results observed for highly proficient bilinguals under slightly different experimental conditions. In Experiment 3, we use a larger set of pictures (40) in order to reduce the number of repetitions of each picture. This modification would also reduce any item specific effects that could have led to the observed difference between L1 and L2 naming latencies.
Experiment 3: Symmetrical switching costs in highly proficient bilinguals

In this experiment we further analyse the switching performance of highly proficient bilinguals. We do so by asking Spanish–Catalan highly proficient bilinguals to perform the same switching task as in the previous experiment. However, in this experiment we include 40 pictures rather than 10. This manipulation allows us to assess whether the pattern of results observed in Experiment 2 generalizes to other experimental contexts (more and different experimental items, and fewer repetitions).

Method

Twelve participants from the same population as in Experiment 2 took part in the experiment. None of them had participated in Experiment 2. Forty pictures with non-cognate names were selected (the 10 pictures used in Experiment 1 plus 30 new pictures) (see Appendix B). The same design and procedure as in the previous experiment was used here. However, unlike in Experiment 2, each picture appeared only once per list, and a given picture was repeated across lists with a minimal interval of five pictures. Each picture appeared 23 or 24 times in the experiment. All the other details were identical to those of Experiment 2.

Results and discussion

We excluded 5.6% of the trials from the analyses (errors: 4.1%; outliers: 1.5%). In the error analysis the main effect of the variable “Type of Trial” was significant ($F(1, 11) = 6.45; \text{MSE} = 1.84; p = .027$; $F2(1, 39) = 4.67; \text{MSE} = 9.99; p = .037$), revealing that participants made more errors on switch (6.3%) than on non-switch (5.3%) trials. Neither the main effect of “Response Language” nor the interaction between “Type of Trial” and “Response Language” were significant (all $F$s < 1).

In the analysis of naming latencies, the main effects of “Type of Trial” ($F1(1, 11) = 30.72; \text{MSE} = 1757.35; p = .001$; $F2(1, 39) = 238.17; \text{MSE} = 686.83; p = .001$) and “Response Language” ($F1(1, 11) = 21.19; \text{MSE} = 4543.17; p = .002$; $F2(1, 39) = 81.13; \text{MSE} = 3817.88; p = .001$) were significant. The interaction between the two variables was not significant (both $F$s < 1). The switching cost for both languages was very similar (L1: 65 ms; L2: 69 ms) (see Fig. 2B).

The results of this experiment fully replicate the effects observed in Experiment 2: (a) to switch from L1 to L2 takes the same amount of time as to switch from L2 to L1; and (b) naming responses in L1 are slower than in L2. Thus, this experiment helps us to extend the observations made in the previous experiment to naming contexts in which the response set is larger and the number of item repetitions is smaller. Also, it reveals that the difference between L1 and L2 naming latencies is unlikely to reflect item specific differences between the L1 and L2 picture names (see Discussion of Experiment 4).

The naming performance of highly proficient bilinguals is consistent with the notion that language switching entails inhibition of the non-response language, and that the amount of inhibition applied to a given language depends on the difference in L1–L2 proficiency. In this framework, asymmetrical switching costs are a function of the different proficiency levels of the two languages involved in the switching task. Consequently, when L2 proficiency is almost as high as that of L1, the asymmetrical switching cost disappears (Experiments 2 and 3). Thus, the present pattern of results is consistent with the notion that inhibitory control plays a role in lexical access in bilingual speakers.

However, the results of Experiment 3 do not necessarily demonstrate the existence of inhibitory control in highly proficient bilinguals. In fact, these results could be interpreted as reflecting either that the two languages of a highly proficient bilingual are inhibited to the same extent, or that the language switching task is achieved without inhibitory processing. That is, it is possible that highly proficient bilinguals have developed a different sort of selection mechanism that does not require inhibition of the non-response language for successful selection of words in the intended language. If that were to be the case, switching from L1 to L2 should be as costly as switching from L2 to L1, which is precisely the observed result.

Thus, the question that remains unanswered is whether highly proficient bilinguals make use of reactive inhibitory mechanisms to perform the switching task. A possible way to answer this question is to test whether highly proficient bilinguals show asymmetrical switching costs when performing the switching task in two languages for which they have different levels of proficiency. To assess this issue we need to test highly proficient bilinguals that are learners of another language (L3) in a switching task involving L1 and L3. If inhibitory control is at the root of the switching performance of highly proficient bilinguals we should expect an asymmetrical switching cost when they perform the task in L1 and L3. In fact, the pattern of results in this case should resemble that observed for L2 learners in Experiment 1.

Experiment 4: Highly proficient bilinguals switching between L1 and L3

The main goal of this experiment is to test whether highly proficient bilinguals show asymmetrical switching costs when asked to perform a switching task in their L1
and in their much weaker L3. We assessed this issue by testing highly proficient Spanish–Catalan bilinguals who were learning English (L3) in a language switching task. Crucially, the L2 (Catalan) proficiency level of these speakers is very different from their L3 (English) proficiency level. In addition, the participants tested in this experiment were selected from the same population as those who participated in Experiments 2 and 3, and therefore they had a similar level of L2 proficiency as in those experiments (see Appendix A).

**Method**

**Participants**

Twelve participants from the same population as that used in Experiment 2 took part in the experiment. All participants had taken English courses for at least 6 years (1 h lessons twice per week), and were currently attending English courses as a second language. They scarcely used English for usual communication. As the language proficiency self-assessment shows, these subjects’ L3 proficiency levels and the proficiency levels of the L2 learners group of the Experiment 1 were very similar. Consequently, participants in this group can be considered to be L3 learners. None of the participants had taken part in the previous experiments.

**Materials and procedure**

The same materials and procedure used in Experiment 1 were employed here. The only difference is that participants were asked to perform the task in their L1–L3 (Spanish–English).

**Results**

We excluded 6.1% of the trials from the analyses (errors: 4.4%; outliers: 1.7%). Participants made slightly more errors on switch (8.2%) than in non-switch trials (6.7%) (“Type of Trial” $F(1, 11) = 3.88; MSE = 6.79; p = .074; F2(1, 9) = 9.89; MSE = 2.29; p = .012$).

In the analysis of naming latencies two main effects were significant: “Type of Trial” ($F(1, 11) = 30.73; MSE = 880.90; p = .001; F2(1, 9) = 105.87; MSE = 223.55; p = .001$), and “Response Language” ($F(1, 11) = 12.84; MSE = 1325.06; p = .004; F2(1, 9) = 12.87; MSE = 1060.58; p = .006$). The interaction between the two variables was not significant (both $F$s < 1). The magnitude of the switching cost was practically identical for the two languages (L1:47 ms; L2: 48 ms) (see Fig. 3).

These results resemble those observed in the same population when performing the task in their two dominant languages L1–L2, in the sense that the magnitude of the switching cost was not modulated by the language of the response. To assess whether the performance of highly proficient bilinguals when performing the task in their L1–L2 or in their L1–L3 is comparable, we ran a joint analysis of the naming latencies obtained for highly proficient bilinguals in Experiment 2, and in the present experiment. The main effects of “Type of Trial” and “Response Language” were significant ($F(1, 22) = 99.71, MSE = 540.21; p = .001; F2(1, 9) = 152.32, MSE = 306.3; p = .001; and $F1(1, 22) = 19.79; MSE = 987.73; p = .001; F2(1, 9) = 10.12, MSE = 1577.13; p = .011$; respectively). Crucially, none of the interactions were significant. Thus, Spanish–Catalan highly proficient bilinguals performed the switching task similarly, regardless of whether it involved their two dominant languages (Spanish–Catalan) or their dominant language (Spanish) and a much weaker language (English–L3). We also explored whether the performance of the bilinguals tested in this experiment differed from that of the L2 learners tested in Experiment 1. Recall that the proficiency level in the non-dominant language used in the experiments for the two groups of participants was comparable.

In this analysis, the three-way interaction between “Type
of Trial”, “Response Language”, and “Experiment” was significant ($F(1, 22) = 5.45; MSE = 368.24; p = .029; $F(2, 9) = 15.14; MSE = 113; p = .004$), indicating that the magnitude of the asymmetrical switching cost was different for the two groups (L2 Learners: 36 ms; L3 learners: -1 ms).²

Another way to assess whether the magnitude of the switching cost in L1 is linked to bilinguals’ proficiency levels in their weaker languages is to explore whether naming latencies in the weaker language predict the difficulty of switching to L1. In this analysis we took the naming latencies in the L2 (or L3) non-switch trials of highly proficient bilinguals in Experiments 2 and 4 as an indicator of participants’ L2 (or L3) proficiency level, and we evaluated whether that variable was a good predictor of the magnitude of the L1 switching cost. The correlation between the two variables was very low (.16) and accounts for a very small part of the variance (.026), supporting the notion that the level of proficiency of the less dominant language involved in the switching task does not affect the ease with which the L1 switch is achieved.

Discussion

The results of this experiment are clear: highly proficient bilinguals do not show asymmetrical switching costs when performing the switching task in their L1 and in their much weaker L3. This result is at odds with the account of the asymmetrical switching cost in terms of an imbalance between the proficiency levels of the two languages of a bilingual, which in turn leads to stronger inhibition of the more dominant language. As one can appreciate in Fig. 4, the difference in proficiency levels between the languages involved in the switching task does not predict the presence of asymmetrical switching costs.

The similarity between the pattern of results in Experiments 2 and 4 also extends to the overall difference in naming latencies between L1 and L3. Surprisingly, in Experiment 4, naming latencies for L1 were slower than for L3, replicating the results of Experiments 2 and 3. Before interpreting the contrasting pattern of results between highly proficient bilinguals and L2 learners regarding the presence (or absence) of asymmetrical switching costs, it is worth discussing the possible origins of this difference between L1, L2, and L3 naming latencies. The first issue we need to assess is the reliability of the phenomenon.

There are two sources of evidence suggesting that the advantage of the non-dominant over the dominant language in language switching tasks is reliable and is not due to specific properties of the materials used in a given experiment. First, to assess whether the difference between L1 and L2 naming latencies stems from poor selection of experimental materials, we asked two different groups of bilinguals, taken from the same population as those tested in Experiment 3, to name the pictures in either their L1 (Spanish) or in their L2 (Catalan). We decided to run this control experiment with the materials used in Experiment 3, because it was in that experiment in which the difference between L1 and L2 was largest. In this control experiment there was no switching task and participants ($n = 24$) only named the pictures in one language (i.e., participants were set in the so-called monolingual mode; Grosjean, 2001): half of the participants named the pictures in their L1

² Following a similar analysis conducted by Meuter and Allport (1999) we explored whether participants’ performance varied throughout the experimental session. In their study, Meuter and Allport found that the asymmetrical switching cost was somewhat reduced in latter parts of the Experiment. Thus, we divided the experiments into two halves, and we declared a new factor “Part of the Experiment”. In Experiment 1, the two-way interaction denoting asymmetrical switching costs (“Type of Trial” × “Response Language”) was present in both halves of the experiment (Group 1: “Part 1”: $F(1, 11) = 15.29; MSE = 341.86; p = .003$; “Part 2”: $F(1, 11) = 6.68; MSE = 235.74; p = .024$; Group 2: “Part 1”: $F(1, 11) = 22.99; MSE = 186.65; p = .001$; “Part 2”: $F(1, 11) = 10.66; MSE = 610.38; p = .008$). The magnitude of the asymmetrical switching cost was similar across the whole Experiment, as revealed by the non-significant interaction between “Type of Trial”, “Response Language”, and “Part of the Experiment” (Group 1: $F(1, 11) = 3.23; MSE = 164.73; p = .10$; Group 2: $F < 1$). In Experiments 2, 3, and 4, participants’ performance was also stable across the Experiment. No interaction was found between the “Type of Trial” and “Response Language” in any part of the experiment for either Experiment (all Fs < 1). Also, the three-way interactions between these two factors and the “Part of the Experiment” were not significant ($F < 1$) in any experiment.
(Spanish) and the other half in their L2 (Catalan). Naming latencies in L1 were faster (606 ms) than in L2 (642 ms). Perhaps, because of the reduced number of participants, this difference only reached significant values in the item analyses ($F(1, 22) = 1.88; MSE = 4721.28; p = .18; F(1, 39) = 50.1; MSE = 563.76; p = .001$). A closer look at the results, revealed that participants named 36 out of the 40 pictures faster in their L1 than in their L2. Thus, it seems that the difference between L1 and L2 is not due to poor selection of the experimental materials (or to poor self-assessment of participants’ language dominance). The second piece of evidence that suggests the reliability of the advantage of L2 over L1, is that such a difference has already been observed in a switching experiment conducted: (a) with another pair of languages (Dutch–English) and; (b) in different laboratories (Janssen, 1999; see also Kroll, Dijkstra, Janssen, & Schriefers, 2000). Thus, given that the phenomenon is reliable, the issue becomes one of understanding its causes.

One possible source of the L2 advantage over L1 is that participants bias the lexicalization process towards their weak language. That is, in some percentage of trials participants may have triggered the lexicalization process in the weak language irrespective of the language cue (the color of the picture). When the language cue is processed according to the proper language schema, and that corresponds to the weak language, the lexicalization process that has already started proceeds with no problems. However, if the language cue corresponds to the dominant language (L1), speakers must stop the lexicalization process in their L2 and start it from scratch in the appropriate language. In this scenario, the advantage of L2 over L1 could stem from: (a) the head start in the lexicalization process of L2, and (b) the “hidden language-switch” involved when naming has to be performed in L1. Note that this account would correspond to the weak language, the lexicalization process that has already started proceeds with no problems. However, if the language cue corresponds to the dominant language (L1), speakers must stop the lexicalization process in their L2 and start it from scratch in the appropriate language. In this scenario, the advantage of L2 over L1 could stem from: (a) the head start in the lexicalization process of L2, and (b) the “hidden language-switch” involved when naming has to be performed in L1. Note that this account would correctly predict, that the difference between L1 and L2/L3 would be independent of the participants’ proficiency in the second language and that such a difference would be present both in switch and non-switch trials. In Experiment 5 we try to shed light on the validity of this explanation.

**Experiment 5: Is the L2 advantage a matter of lexicalization bias?**

In this experiment, we assess whether the difference in naming latencies between the dominant and the non-dominant languages is due to a bias in the naming process that prioritizes the start of lexicalization in the non-dominant language. One property of the design used in the previous experiments may have been critical for establishing such a bias: the simultaneous presentation of the language cue and the target picture. Given this simultaneous presentation, participants may have started encoding the name of the picture and the language cue simultaneously, favoring lexicalization in the less dominant language. One way to discourage such a bias is to reduce the uncertainty regarding the language in which a given picture will be named. If participants know in advance the language in which the picture has to be named there would be no need to bias selection towards the non-dominant language. Thus, if naming latencies in the non-dominant language are faster than in the dominant language because of the bias in the lexicalization process, then a reduction of such a bias by informing participants about the language in which a picture has to be named would result in a reduction of the difference in naming latencies between L1 and L2. Furthermore, as other studies have shown, reducing the uncertainty in a switching task also reduces the magnitude of the switching cost (De Jong, 1995; Meiran, 1996; Rogers & Monsell, 1995).

We test this hypothesis by presenting a cue that indicates the language in which a given picture has to be named before the actual presentation of the picture: each picture is preceded by a colored circle indicating the language in which the picture has to be produced (see Kroll et al., 2000 for a similar design). We use two different stimulus onset asynchronies (SOA) between the presentation of the language cue and the target picture. Participants saw the language cue either 500 ms (Group 1) or 800 ms (Group 2) before the target picture.

**Method**

Twenty-four participants from the same population as in Experiment 2 were randomly assigned to two groups, corresponding to the two SOAs. The same materials used in Experiment 2 were employed here. The procedure was very similar to that of Experiment 2. There were, however, two main differences. First, each picture was preceded by the presentation of a language cue (a red or blue circle) that was displayed for 300 ms. Second, for Group 1, the picture appeared 500 ms after the language cue onset, while for Group 2, each picture appeared 800 ms after the language cue onset.

**Data analysis**

Given that the crucial issue addressed here is whether advancing the presentation of language cue eliminates the difference between naming latencies for L1 and L2, we compare the results obtained when the language cue is presented before the picture (SOA: 500, SOA: 800) with those obtained when both stimuli were presented simultaneously (Experiment 2; SOA: 0). Thus, in the following analyses we declare three variables “Type of Trial” (Switch vs Non Switch), “Response Language” (L1 vs L2), and SOA (SOA 0 (results from Experiment 2), SOA 500, SOA 800).
Results

Following the same criteria as in the previous experiments, 4.9% (errors: 3.4%; outliers: 1.5%) and 5.5% (errors: 3.9%; outliers: 1.6%) of the data points were excluded from the analyses for SOA 500 and SOA 800, respectively.

In the analysis of error rates, the only significant effect was that of “Type of Trial” (F(1, 133) = 21.05; MSE = 2.37; p = .001; F2(1, 9) = 16.62; MSE = 2.39; p = .003), reflecting the presence of more errors on switch trials (5.8%) than on non-switch trials (4.7%).

In the analysis of naming latencies two main effects were significant: “Type of Trial” (F(1, 133) = 53.55; MSE = 628.79; p = .001; F2(1, 9) = 105.36; MSE = 273.63; p = .001), and “Response Language” (F(1, 133) = 10.33; MSE = 1291.92; p = .003; F2(1, 9) = 4.2; MSE = 2618.2; p = .07). Importantly, the magnitude of the switching cost was modulated by the SOA, as reflected by the significant interaction between the variables “Type of Trial” and “SOA” (F(1, 133) = 4.43; MSE = 2790.08; p = .020; F2(1, 9) = 30.67; MSE = 140.88; p = .001). Despite this modulation, the difference in naming latencies between L1 and L2 remained stable across SOAs (SOA 0: 19; SOA 500: 20 ms; SOA 800: 19 ms) as revealed by the non-significant interaction between “Response Language” and “SOA” (both F’s < 1). None of the remaining interactions were significant (Fig. 5).

A further analysis of the interaction between “Type of Trial” and “SOA” revealed that the switching cost (“Type of Trial”) was significant for the three SOAs (SOA 0 (from Experiment 2): F(1, 11) = 134.27; MSE = 199.52; p = .001; F2(1, 9) = 132.37; MSE = 173.69; p = .001; SOA 500: F(1, 11) = 9.88; MSE = 896.19; p = .009; F2(1, 9) = 26.76; MSE = 262.38; p = .001; and SOA 800: F(1, 11) = 4.56; MSE = 790.65; p = .05; F2(1, 9) = 44.1; MSE = 78.02; p = .001). However, the magnitude of the switching cost decreased as the SOA increased (SOA 0: 47; SOA 500: 27; SOA 800: 17). The interaction between “Response Language” and “Type of Trial” was not present at any SOA.

Discussion

The goal of this experiment was to assess one account of the difference in naming latencies between L1 and L2 observed in Experiments 2, 3, and 4. We argued that such a difference may be reflecting a bias to start lexicalization in L2, and that one way to reduce such a bias, and hence such a difference in naming latencies, is to inform participants about the language in which the picture will be named some time before the presentation of the target picture.

The results of the experiment did not support this prediction. Instead, the difference between L1 and L2 naming latencies was unaffected by the time given to participants to prepare the response language. Importantly, the failure to modulate this difference cannot be attributed to participants ignoring the language cue, since the magnitude of the overall switching cost was diminished when the language cue was presented before the target picture (see Yeung & Monsell, 2003, for a discussion of this issue). Thus, it appears that a bias in the triggering of the lexicalization process is not at the basis of the L2 advantage in the language-switching task. In the General Discussion we put forward a tentative explanation of this difference in the context of the other effects reported in this article.

General discussion

We reported five experiments in which the performance of different populations of bilingual speakers in a language switching task was tested. The main issue assessed was whether a difference in proficiency levels between the two languages involved in the switching task affects the switching performance of highly proficient bilinguals and L2 learners to the same extent.
Given that the asymmetrical switching costs observed in this paradigm (switching to the more dominant language is more difficult than switching to the less dominant language) have been taken as evidence supporting the notion that bilingual speech production entails inhibitory control, it is important to assess the presence of such a pattern in different populations of bilingual speakers.

In Experiment 1, two different groups of L2 learners (Spanish–Catalan and Korean–Spanish bilinguals) were asked to perform the switching task in their L1 and in their much weaker L2. In both groups, switching from L2 to L1 was harder than vice versa, replicating Meuter and Allport's (1999) observation.

In Experiment 2, highly proficient Spanish–Catalan bilinguals performed the same switching task as in Experiment 1. However, for these bilinguals the magnitude of the switching cost was the same for L1 and L2, and naming latencies were faster in L2 than in L1. The same pattern of results was observed under experimental conditions that minimized the possible contribution of item specific effects (Experiment 3).

Experiment 4 was designed to further test the hypothesis of inhibitory control in bilingual lexical access. Highly proficient Spanish–Catalan bilinguals who were learners of English were asked to perform the switching task in L1 and in their much weaker L3 (English). The results resembled those observed for highly proficient bilinguals in previous experiments: the magnitude of the switching cost was similar for both languages (Spanish and English).

Finally, in Experiment 5 we evaluated one explanation of the origin of the advantage in naming latencies for the less dominant languages (L2 and L3 faster than L1). We did so by minimizing the chances that participants develop a lexicalization bias for their weaker language. The results of this experiment showed: (a) a reduction of the overall magnitude of the switching cost, (b) no modulation of the L2 advantage over the L1. These results can be summarized as follows:

1. Language switching costs are present in all types of bilingual speakers tested up to now.
2. Asymmetrical switching costs are present for L2 learners (to switch into the dominant language is harder than to switch into the weaker language), but not for highly proficient bilinguals.
3. The switching performance of highly proficient bilinguals is independent of the difference in proficiency levels between the two languages involved in the task.
4. In a language switching task, highly proficient bilinguals are slower in their dominant than in their non-dominant language both on switch and non-switch trials.

Before addressing the theoretical implications of these results, it is important to recall the two mechanisms that have been invoked to explain the origin of language switching costs. The first refers to the change in the task-goal involved in switching trials. In a switching task, participants have to choose between two different tasks on every single trial (e.g., name in language X or name in language Y). According to theories of task switching, to change the goal of the task (or the language schema) from one trial to another (switching trials) takes more time than re-using the same one (non-switch trials). The switching cost can thus be assumed to reflect the time needed to re-set the cognitive system and select a new task-goal, a time cost that will be present regardless of the ease of the tasks involved in the switching task. Thus, bilingual speakers would experience an overall switching cost when changing the “language schema” (e.g., when switching), regardless of their language proficiency (see also Kohnert, Bates, & Hernandez, 1999; for effects of switching languages across the life-span of bilingual speakers). The second component refers to the differential inhibition applied to the lexical representations of the languages involved in the switching task. According to some researchers, the successful selection of a given lexical representation in the proper language requires the reactive inhibition of lexical representations in the other language. In Green's proposal (e.g., Green, 1998) this inhibitory mechanism is sensitive to the different levels of activation of the two languages of a bilingual, in such a way that the amount of inhibition applied to the more available (i.e., dominant) language would be larger than that applied to the less available language. This differential amount of inhibition and its effects on subsequent trials would account for the existence of asymmetrical switching costs when the two languages involved in the switching task differ in their strength (availability). Let us now evaluate the theoretical implications of the four experimental observations reported in this article.

The first main result of this study is that all bilingual speakers showed language switching costs. This effect is predicted by the first mechanism involved in the language switching task: the extra time needed to reset the cognitive system and to choose a different language schema from that which has been used in the previous trial. Note, that the presence of this effect is independent of the precise nature of the lexical selection mechanism. That is, regardless of whether inhibitory control is postulated, the change in a “language schema” will always have an associated time cost.

Of special relevance for our interests are the observations summarized in points 2 and 3 above. These results allow us to draw the following empirical generalization: The crucial factor for the presence of asymmetrical switching costs is whether or not participants are highly proficient bilinguals in any given pair of
languages. That is, a difference in dominance between the languages involved in the switching task results in asymmetrical switching cost for L2 learners but not for highly proficient bilinguals. The existence (or lack) of asymmetrical switching costs has often been considered as a signature effect of the presence of inhibitory process. The fact that highly proficient bilinguals do not show asymmetrical switching costs when the switching task involves the two languages for which they are highly proficient is consistent with this interpretation (similar levels of proficiency, similar levels of inhibition). Problematic for this account, however, is the observation that highly proficient bilinguals do not show asymmetrical switching costs when the task involves their L1 and L3. If the relevant factor for observing asymmetrical switching costs is the difference in the amount of inhibition applied to L1 and L2 (as in the case of L2 learners), then such an asymmetry should have been present when highly proficient bilinguals perform the task in their strong L1 (Spanish) and much weaker L3 (English). Thus, this observation casts some doubts on whether the explanation given to the switching performance of L2 learners can account for the performance of highly proficient bilinguals. More importantly, the issue arises of why a difference in the proficiency levels of the two languages involved in the switching task leads to asymmetrical switching costs for L2 learners but not for highly proficient bilinguals.

Why is the switching performance of highly proficient bilinguals and L2 learners different?

The difference in performance between highly proficient bilinguals and L2 learners could be explained in terms of the different degrees with which these two populations master the switching task. It is possible that, by virtue of having practiced language-switching more often, the inhibitory control system of highly proficient bilinguals works in a different way from that of L2 learners. For instance, one could argue that while the inhibitory system of L2 learners takes into account the different strengths of the languages involved in the switching task, such a variable is irrelevant for the inhibitory system of the highly proficient bilinguals. That is, while L2 learners would inhibit their L1 more than their L2, highly proficient Spanish–Catalan bilinguals/learners of English would inhibit all of their languages to the same degree, irrespective of the strength (or availability) of the specific languages involved in the switching task. On this view, to become a highly proficient bilingual would, among other things, entail the development of an inhibitory mechanism that is not reactive; that is, a mechanism that does not depend on the relative strength of the two languages involved in the switching task (but see below for problems found by this view when trying to account for the difference in naming latencies between L1/L2).

However, given the striking difference between the switching performance of highly proficient bilinguals and L2 learners, one could seek an explanation in terms of different lexical selection mechanisms between the two populations rather than in terms of the degree with which they inhibit their different languages. That is, it is possible that the difference between the ways in which these two populations perform the switching task is not quantitative (e.g., amount of inhibition applied to the languages) but qualitative (e.g., the way in which lexical selection is achieved). As advanced in the Introduction, there are at least two views about how lexical selection proceeds in bilingual speakers. According to one of these explanations, lexical selection in the intended language is achieved by means of the active suppression of the lexical items belonging to the non-response language. Alternatively, the language-specific selection hypothesis assumes that lexical access does not require inhibitory control. In this framework, a bilingual speaker has developed a lexical selection mechanism that is only sensitive to the activation levels of the words belonging to the intended language. Thus, the level of activation of the words belonging to the non-response language would not act as competitors at the level at which lexical selection is achieved, and therefore, there will be no need to suppress their activation.

In this scenario, one could put forward the following tentative explanation: the degree with which lexical selection in bilingual speakers entails inhibitory control depends on whether they have achieved a high proficiency level in any pair of languages. If the bilingual has not achieved comparable performance levels in her two languages, then lexical selection would make use of inhibitory control to ensure selection in the intended language. However, an increase in the proficiency level of the bilingual speaker would lead to a shift in the “type” of processes responsible for focusing on one language. The shift would be from the reliance on inhibitory control to the reliance on a language-specific selection mechanism during lexical selection. Importantly, when this specific selection mechanism has been developed, it will be applied to any language, regardless of the proficiency level of the speaker in that language. That is, the specific selection mechanism would be functional in the dominant L2, and in the weaker L3. In other words, if speakers are able to focus their lexical selection in the strong L2 without suppressing L1, they will also be able to do so in the much weaker L3. On this view, given that reactive inhibition is only present for L2 learners, the asymmetrical switching cost would be present only for this group of speakers. And, highly proficient bilinguals would not show such an asymmetrical switching cost regardless of the difference in proficiency levels between the languages involved in the switching task (although
they will suffer overall switching costs, as discussed above).3

The notion that bilingualism may exert an influence on the way a third language is processed has already been put forward by several authors (e.g., Cenoz & Valencia, 1994; Lasagabaster, 2000; Sanz, 2000). Some of these studies reveal that even when several important variables such as general intelligence or similarity between languages are controlled for, bilingual speakers show an advantage in learning an L3 (see Cenoz, 2003, for a review). Furthermore, differences in the acquisition of an L3 seem to correlate with the degree of bilingual proficiency. For example, the study of Sagasta (2003) showed that a higher level of bilingualism in Spanish–Basque bilinguals correlated with a better performance in acquiring English as an L3.

Up to here we have discussed the results summarized in points 1–3, and we argued that it is possible to account for: (a) the presence of switching costs in all types of bilinguals, (b) the presence of asymmetrical switching costs for L2 learners, and (c) the presence of symmetrical switching costs for highly proficient bilinguals. In the following we address the implications of the fourth observation: highly proficient bilinguals named the pictures in L1 more slowly than in L2 or L3.

Accounting for the performance of highly proficient bilinguals: The advantage of L2 (L3) over L1 and the symmetrical switching costs

In Experiment 5, we attempted to clarify the origin of the L2 advantage over L1. We assessed whether a bias to start lexicalization in the less dominant language, regardless of the language cue, was at the basis of the L2 advantage over L1. The results of Experiment 5 failed to support such a hypothesis, and therefore we rule out such a bias as the critical factor behind the difference in L1–L2 naming latencies.

At first glance, the advantage of L2 (or L3) over L1 could be taken as revealing the existence of larger inhibition for L1 lexical representations in the switching task. In other words, naming latencies in L1 are slower than in L2 (or L3) because the L1 representations are more suppressed than those belonging to the other languages. Although this account captures the L1 disadvantage, it fails to explain another crucial result in our investigation: the lack of asymmetrical switching costs. Thus, if in order to explain the L1 disadvantage in naming latencies we assume that L1 is more inhibited than L2 (L3), then we are compelled to incorrectly predict the presence of asymmetrical switching costs. This is because the larger the inhibition applied to a given language the larger the switching cost associated with it.

Thus, it appears that the assumption of inhibition creates problems when accounting for both the L2(L3) advantage over L1 and the presence of symmetrical switching costs. How can we reconcile these two observations? Perhaps a better solution is to dispense with the assumption of inhibitory control in highly proficient bilinguals. In the following we entertain one possible explanation that we refer as to the “language-specific selection threshold hypothesis.”

One of the factors determining the speed of picture naming is how fast lexical selection is achieved. At the same time, there is wide agreement in assuming that the speed of lexical selection depends, among other things, on the availability of lexical representations. For example, the lexical nodes corresponding to pictures with high frequency names are selected and produced faster than those with low frequency names (Alario, Costa, & Caramazza, 2002; Caramazza, Costa, Miozzo, & Bi, 2001; Jescheniak & Levelt, 1994; Oldfield & Wingfield, 1965). In this scenario, one should expect L1 responses to be faster than L2 or L3 responses (at least in language-blocked naming conditions), given that L1 representations are presumably more available than the others (see for example, the control experiment presented in the discussion section of Experiment 4). However, in the context of a language switching task, participants may have tried to compensate such an imbalance by making the lexical representations of the weaker language more available. There are several ways in which such an increase in the availability of L2/L3 lexical representations may have come about. For example, it is possible that the criteria for lexical selection to proceed (e.g., the level of activation of the target in relation to other lexical nodes) can be set independently for different languages (language-specific). In this way, the time required for an L2 word to reach the selection criterion would be shorter than for an L1 word. This imbalance will in turn lead to faster selection and production of L2 words. According to this hypothesis, speakers should be able to use information about the task.

3 The explanations given for the difference in performance between highly proficient bilinguals and L2 learners retain the notion that the asymmetrical switching cost indexes the presence of inhibitory process. However, one could try to account for such an effect without resorting to this assumption. It is possible that the asymmetrical switching costs reflects the differential amount of general resources consumed by naming in L1 and naming in L2. That is, if we assume that to perform a naming task in L2 is harder than in L1, then when the subsequent trial requires a switch into L1, there would be relatively few resources left for the cognitive system to change the “language schema”. On the other hand, when the task requires to switch into L2, the cognitive system would have relatively more resources to change the “language schema”. This imbalance will translate into a larger switching cost for L1 than for L2. However, embracing this explanation of the asymmetrical switching costs does not answer the question of why highly proficient bilinguals do not show such an asymmetry when the languages involved in the switching task would presumably consume different amount of resources (i.e., Experiment 4).
We have interpreted the results presented in this article focusing on the relative difference in the speakers’ proficiency levels. We have argued that high proficiency levels in L2 may have important consequences not only on how language switching is performed in L1 and L2, but also in L1 and in the less strong L3. However, as advanced in the Introduction of this article, there is an alternative explanation for these results, in which the relevant variable is not L2 proficiency but L2 age of acquisition. That is, our highly proficient bilinguals may behave differently from the less-proficient bilinguals not because their L2 proficiency level is higher but rather because they acquire their L2 earlier in life. The results reported in this article do not allow us to tease apart the independent contributions of these two variables. This is because the highly proficient bilinguals tested in our experiments had a high proficiency level in L2 and had acquired that language very early in life. However, a tentative answer to this issue can be found when we compare our results to those reported by Hernandez and Kohnert (1999). In their study, Spanish–English bilinguals that learned their L2 before the age of 8 were tested. However, and despite Spanish being the first language of the participants, their dominant language was English. For example, in a blocked naming condition responses in English were faster than in Spanish. Thus, because of their extensive and intensive exposure to English their second language became the more dominant one, revealing a dissociation between L1/L2 relative dominance. The results of this experiment showed a larger switching cost in the dominant language (English) than in the less dominant language (Spanish), replicating the asymmetrical switching cost observed in Experiment 1. If we compare these results with those observed for the Spanish–Catalan bilinguals tested in Experiments 2 and 4, we might tentatively conclude that proficiency rather than age of L2 acquisition is at the basis of the different performance between the bilingual groups. This is because both groups behave differently despite the fact that they were exposed to their two languages early in life. However, future research needs to be conducted to evaluate more directly the contributions of age of acquisition and proficiency levels to the observed pattern of results.

**Conclusion**

The experiments reported in this article revealed striking differences in the language-switching performance of L2 learners and highly proficient bilinguals. While the former group showed a clear asymmetrical
switching cost (e.g., switching to L1 being harder than switching to L2), the latter did not, even when the difference in proficiency levels between the two languages tested in the experiment was large. Furthermore, highly proficient bilinguals were faster when naming the pictures in their less dominant language than in their dominant language. We argued that it is difficult to find a single unitary explanation for the performance of both groups of bilinguals in terms of the reactive inhibitory control proposed by Meuter and Allport (1999). We further argued that such a difference may be revealing differences in the types of mechanisms involved in lexical access for low and highly proficient bilinguals.

Furthermore, and regardless of the precise origin of the different performance profiles of highly proficient bilinguals and L2 learners, our results reveal that becoming a highly proficient bilingual seems to have implications not only for the way the two dominant languages (L1–L2) are processed, but also for the way a much weaker language is processed (L3). That is, the effects of bilingualism on the way lexical access is achieved are not limited to the two languages for which the bilingual is proficient, but they extend to other linguistic contexts, independent of the degree with which those contexts are mastered by the speaker (see Cenoz, 2003, for a review on third language acquisition by bilingual and monolingual speakers). It is an open question whether highly proficient bilingualism may affect performance in other switching tasks that do not involve linguistic components (see Bialystok, 1999, in press).

Acknowledgments

This research was supported by two grants from the Spanish Government (BSO2001-3492-C04-01; BFF2002-10379-E), and by the McDonnell grant “Bridging Mind Brain and Behavior.” Albert Costa was supported by the research program “Ramon y Cajal” from the Spanish government. Mikel Santesteban was supported by a Predoctoral fellowship from the Basque government (PI-1999-18). Requests for reprints may be addressed to Albert Costa. The authors are grateful to Gabi Felhosi and Jin Nam Choi for their assistance in running the experiments, and to Nuria Sebastián-Gallés, Alfonso Caramazza, Saamah Abdallah, Salvador Soto, Antoni Rodriguez, Matthew Finkbeiner, Mark Knobel, Bradford Mahon, Niels Janssen, and Laura Bosch for their comments.

Appendix A. Language history and the self-assessed proficiency for all participants

Language history and the self-assessed proficiency scores of the participants in all the experiments reported in the article.

Appendix B. Materials employed in all the experiments

Materials used in the experiments. In Experiment 3, the whole set of words were used. In Experiments 1, 2, 4, and 5 only the 10 words marked with an asterisk were used. The materials marked with the symbol @ were used in the Korean-Spanish experiment. The Catalan and Spanish names of the pictures were matched for number of phonemes (5.7 vs 5.6, respectively). Each word and its translation had the same gender value. The Korean and Spanish names were matched for number of phonemes, 4.4 and 5.6, respectively (r(9) = 1.42, p = .187). In Experiment 3, care was taken to select a collection of pictures from various semantic categories and to select those pictures for which their corresponding Spanish and Catalan words were of...
similar length: 5.1 vs 5.2 number of phonemes, respectively $(r(39) = .52, p = .005)$.

<table>
<thead>
<tr>
<th>Spanish name</th>
<th>Catalan name</th>
<th>English name</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Sombrero @</td>
<td>Barret</td>
<td>Hat</td>
</tr>
<tr>
<td>*Zanahoria @</td>
<td>Pastanaga</td>
<td>Carrot</td>
</tr>
<tr>
<td>*Manzana @</td>
<td>Poma</td>
<td>Apple</td>
</tr>
<tr>
<td>*Mesa</td>
<td>Taula</td>
<td>Table</td>
</tr>
<tr>
<td>*Perro @</td>
<td>Gos</td>
<td>Dog</td>
</tr>
<tr>
<td>*Hoja @</td>
<td>Fulla</td>
<td>Leaf</td>
</tr>
<tr>
<td>*Cuchillo @</td>
<td>Ganiyet</td>
<td>Knife</td>
</tr>
<tr>
<td>*Ventana @</td>
<td>Finestra</td>
<td>Window</td>
</tr>
<tr>
<td>*Lluvia @</td>
<td>Pluja</td>
<td>Rain</td>
</tr>
<tr>
<td>*Queso</td>
<td>Formatge</td>
<td>Cheese</td>
</tr>
<tr>
<td>Rama</td>
<td>Branca</td>
<td>Branch</td>
</tr>
<tr>
<td>Gusano</td>
<td>Cuc</td>
<td>Worm</td>
</tr>
<tr>
<td>Red</td>
<td>Xarxa</td>
<td>Net</td>
</tr>
<tr>
<td>Jaula</td>
<td>Gàbia</td>
<td>Cage</td>
</tr>
<tr>
<td>Corcho</td>
<td>Suro</td>
<td>Cork</td>
</tr>
<tr>
<td>Jamón</td>
<td>Pernil</td>
<td>Ham</td>
</tr>
<tr>
<td>Melocotón</td>
<td>Pressec</td>
<td>Peach</td>
</tr>
<tr>
<td>Mujer</td>
<td>Dona</td>
<td>Woman</td>
</tr>
<tr>
<td>Hueso</td>
<td>Os</td>
<td>Bone</td>
</tr>
<tr>
<td>Dedo</td>
<td>Dit</td>
<td>Finger</td>
</tr>
<tr>
<td>Huevo @</td>
<td>Ou</td>
<td>Egg</td>
</tr>
<tr>
<td>Ojo</td>
<td>Ull</td>
<td>Eye</td>
</tr>
<tr>
<td>Pañuelo</td>
<td>Mocador</td>
<td>Handkerchief</td>
</tr>
<tr>
<td>Muleta</td>
<td>Crossa</td>
<td>Crutch</td>
</tr>
<tr>
<td>Calçetín</td>
<td>Mitjó</td>
<td>Socket</td>
</tr>
<tr>
<td>Rana</td>
<td>Granota</td>
<td>Frog</td>
</tr>
<tr>
<td>Pimiento</td>
<td>Pobrot</td>
<td>Pepper</td>
</tr>
<tr>
<td>Vela</td>
<td>Espelma</td>
<td>Candle</td>
</tr>
<tr>
<td>Búho</td>
<td>Mussol</td>
<td>Owl</td>
</tr>
<tr>
<td>Cepillo</td>
<td>Raspall</td>
<td>Brush</td>
</tr>
<tr>
<td>Cerdo</td>
<td>Porc</td>
<td>Pig</td>
</tr>
<tr>
<td>Silbato</td>
<td>Xiuilet</td>
<td>Whistle</td>
</tr>
<tr>
<td>Pato</td>
<td>Anec</td>
<td>Duck</td>
</tr>
<tr>
<td>Mariposa</td>
<td>Papallona</td>
<td>Butterfly</td>
</tr>
<tr>
<td>Silla @</td>
<td>Cadira</td>
<td>Chair</td>
</tr>
<tr>
<td>Mancha</td>
<td>Taca</td>
<td>Blot</td>
</tr>
<tr>
<td>Burro</td>
<td>Ase</td>
<td>Donkey</td>
</tr>
<tr>
<td>Hoz</td>
<td>Falç</td>
<td>Sickle</td>
</tr>
<tr>
<td>Zueco</td>
<td>Esclop</td>
<td>Clog</td>
</tr>
<tr>
<td>Colchón</td>
<td>Matalàs</td>
<td>Mattress</td>
</tr>
</tbody>
</table>

References


