How Do Highly Proficient Bilinguals Control Their Lexicalization Process? Inhibitory and Language-Specific Selection Mechanisms Are Both Functional

Albert Costa, Mikel Santesteban, and Iva Ivanova
Grup de Recerca en Neurociència Cognitiva, Parc Científic de Barcelona, and Universitat de Barcelona

The authors report 4 experiments exploring the language-switching performance of highly proficient bilinguals in a picture-naming task. In Experiment 1, they tested the impact of language similarity and age of 2nd language acquisition on the language-switching performance of highly proficient bilinguals. Experiments 2, 3, and 4 assessed the performance of highly proficient bilinguals in language-switching contexts involving (a) the 2nd language (L2) and the L3 of the bilinguals, (b) the L3 and the L4, and (c) the L1 and a recently learned new language. Highly proficient bilinguals showed symmetrical switching costs regardless of the age at which the L2 was learned and of the similarities of the 2 languages and asymmetrical switching costs when 1 of the languages involved in the switching task was very weak (an L4 or a recently learned language). The theoretical implications of these results for the attentional mechanisms used by highly proficient bilinguals to control their lexicalization process are discussed.

Keywords: bilingualism, language production, lexical access, language switching

Despite the ease with which language production is conducted, the cognitive machinery responsible for this ability is complex: It involves the integration of pragmatic, semantic, syntactic, phonological, and articulatory processes and representations (e.g., Caramazza, 1997; Dell, 1986; Garrett, 1980; Levelt, 1989). The attentional/control mechanisms in charge of orchestrating all these processes face an additional burden in cases of bilingualism: A bilingual person has to ensure that the intended message is conveyed in the proper language for the communicative context. Thus, a fundamental question in bilingual production is: What is the control mechanism guaranteeing output in the desired language? Despite the growing interest in this issue (e.g., Costa, Miozzo, & Caramazza, 1999; Costa & Santesteban, 2004a; de Bot, 1992; Gollan & Acenas, 2004; Gollan & Kroll, 2001; Hartsuiker, Pickering, & Veltkamp, 2004; Hermans, Bongaerts, de Bot, & Schreuder, 1998; Lee & Williams, 2001; Poullise & Bongaerts, 1994; for reviews, see Costa, 2005; La Heij, 2005), a detailed explanation is still lacking. In the present article, we focus on the nature of the language control mechanism in highly proficient bilinguals.

Current models of bilingual production assume that in the course of lexicalization the semantic system activates in parallel the two languages of a bilingual person (Colomé, 2001; Costa, Caramazza, & Sebastián-Gallés, 2000; Costa, Santesteban, & Caño, 2005; Green, 1986, 1998; Hermans et al., 1998; La Heij, 2005; Roelofs, 1998; but see Costa, La Heij, & Navarrete, 2006, for a critical discussion of the evidence supporting this notion). Given this parallel activation, how does the speaker guarantee that the correct word rather than its translation is selected?

For the so-called inhibitory models, selection is achieved by an inhibitory mechanism that suppresses the activation of the lexical representations of the nonresponse language (e.g., Green, 1986, 1998; Lee & Williams, 2001; Meuter & Allport, 1999). Suppression is applied only after a lexical node has been activated from the conceptual system (therefore the mechanism is assumed to be reactive) and is proportional to the level of activation of a particular item (i.e., the more strongly activated the item, the more inhibition is needed to prevent it from being produced). Thus, when speaking in the first language (L1), not much inhibition is required to suppress 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 items. However, when speaking in L2, L1 representations must be strongly inhibited to ensure selection of L2 lexical items.

Other models, however, do not assume inhibitory mechanisms but rather postulate that the lexical selection mechanism is sensitive to the language membership of lexical representations and considers for selection only those belonging to the response language—the so-called language-specific selection hypothesis (e.g., Costa & Caramazza, 1999; Costa et al., 1999; Roelofs, 1998).
According to this view, lexical representations of the nonresponse language do not enter into competition during lexical selection, and lexicalization proceeds similarly to that in monolinguals (see Costa, 2005, for a detailed description of this mechanism). This view can be implemented in different ways. For example, in Roelofs’s (1998) proposal, target selection is achieved by means of a set of production rules that operate only when some conditions are met (action–condition pairs, if–then rules). One of those conditions can be language membership; if such a condition is not met, selection does not take place. The idea that lexical selection is sensitive to lexical properties was present already in Dell’s (1986) model, in which only those words corresponding to the grammatical category of the intended word are considered for selection. If we extrapolate such a grammatically constrained selection mechanism to bilingualism, we only need to further assume that lexical representations are integrated into a lexicon and marked for the language to which they belong. As we will see below, despite the similarities between the language-specific selection model and the inhibitory model, they make different experimental predictions.

Most of the evidence about the nature of bilingual lexical selection comes from two paradigms: the picture–word interference paradigm (and its variants; e.g., Costa, Colomé, Gómez, & Sebastián-Gallés, 2003; Costa et al., 1999; Ehri & Ryan, 1980; Goodman, Haith, Guttentag, & Rao, 1985; Mägiste, 1984, 1985; Miller & Kroll, 2002; Smith & Kirnser, 1982) and the language-switching paradigm (Costa & Santesteban, 2004b; Hernandez, Dapretto, Mazzotta, & Bookheimer, 2001; Hernandez & Kohnert, 1999; Hernandez, Martínez, & Kohnert, 2000; Jackson, Swainsone, Cunnington, & Jackson, 2001; Meuter & Allport, 1999; Von Studnitz & Green, 2002). Of special relevance in the present context is the information gathered with the language-switching task. In this task, participants are asked to name pictures (or other stimuli, such as Arabic numbers), alternating between their two languages. The language in which a given stimulus has to be named is indicated by a cue (e.g., L1 if the color is blue and L2 if it is red). There are trials in which the stimulus is named in the same language as that used on the preceding trial (nonswitch trials) and trials in which the stimulus is named in a different language than that used on the previous trial (switch trials). The difference in naming latencies between switch and nonswitch trials is referred to as the language-switching cost.

Using this task, Meuter and Allport (1999) reported a seemingly paradoxical result: L2 learners suffered a larger switching cost in their L1 than in their L2. This asymmetrical switching cost was interpreted as supporting the notion of reactive inhibition of the lexical representations belonging to the nonresponse language. The logic behind this conclusion is the following: Switching into the dominant language (switch into L1) is more difficult because its lexical representations have been strongly inhibited on the previous trial to allow the selection of L2 lexical representations. Thus, part of the switching cost is due to the time required to overcome this inhibition. Switching into the weak language (switch into L2) is relatively easier because, when naming in L1, its corresponding lexical representation would not have been inhibited that strongly, and therefore the L2 representations would be more available. In short, switching into L1 is harder than switching into L2 because the L1 is inhibited to a greater extent than the L2. This interpretation of the asymmetrical switching cost makes a straightforward prediction: Highly proficient bilinguals should show symmetrical switching costs. This is because the smaller the difference in language proficiency, the smaller the difference in the level of inhibition applied to the two languages. Costa and Santesteban (2004b) proved this prediction to be right. In our experiments, L2 learners (Spanish learners of Catalan, and Korean learners of Spanish) showed the predicted asymmetrical switching cost. More important, the same switching cost for L1 and L2 was observed for highly proficient Spanish–Catalan bilinguals.

These results lead us to the following empirical generalization: Highly proficient bilinguals do not show asymmetrical switching costs but low-proficient bilinguals do. By putting the emphasis on the proficiency level of bilinguals, one main assumption of the inhibitory account seems to be supported: The amount of inhibition applied to the nonresponse language depends on the difference between the proficiency levels of the two languages of a bilingual. However, this conclusion is premature because the highly proficient bilinguals we tested not only were proficient in their L2 but also had acquired it very early in their lives. Furthermore, they were bilinguals of two lexically and grammatically very similar Romance languages, Catalan and Spanish (75% of the words are cognates). Thus, one could capture the present set of results without appealing to proficiency but rather to other variables: Symmetrical switching costs will only be obtained for those bilinguals who have acquired a high proficiency level in two very similar languages or who have acquired their L2 very early in their lives. To determine which one of these characterizations is correct is of theoretical relevance. This is because, if other populations of highly proficient bilinguals do not show symmetrical switching costs, then the interpretation of the symmetrical pattern of results in terms of equal amounts of inhibition should be reconsidered. And, consequently, we could not use the results from highly proficient bilinguals to support the inhibitory model. Experiment 1 aims to assess this issue.

Age of Second Language Acquisition and Language Similarity

The age at which L2 is acquired has important and pervasive effects in many bilingual domains (Birdsong, 1999; Birdsong & Molis, 2001; Flege, 1999; Flege, Yeni-Komshian, & Liu, 1999; Newport, 1990, 1991; Sebastián-Gallés & Bosch, 2002; Sebastián-Gallés, Echeverría, & Bosch, 2005; Sebastián-Gallés & Soto-Faraco, 1999; Weber-Fox & Neville, 1996, 1999; for a review, see Sebastián-Gallés & Kroll, 2003). This variable affects not only the acquisition of representations, such as phonological representations, but also the mastering of processing mechanisms, such as syntactic and morphological processing. In fact, Wartenburger et al. (2003) argued that different linguistic components are affected by proficiency and age of acquisition to different extents: Profi-
iciency seems to play a greater role for semantic information, whereas age of acquisition affects grammatical knowledge. Furthermore, early acquisition of an L2 has consequences outside of the linguistic domain (as in the development of theory of mind; Goetz, 2003). Of special interest are the benefits that early bilingualism exerts in the development of executive control functions in children and in adults (Bialystok, 1999, 2005). Given these considerations, it is appropriate to explore the role of L2 age of acquisition in the attentional mechanisms involved in bilingual production.

Language similarity may also affect the way bilinguals control their speech. First, there are reasons to believe that the more similar the two languages of a bilingual are, the more likely it is they interfere with each other. In some sense, bilingual lexical access is a situation in which conflict between two potential responses needs to be resolved (bilinguals need to attend to some representations while ignoring others). The ease with which such a conflict is resolved may depend, to some extent, on the similarity between the ignored and attended representations. The greater the similarity between them, the more difficult conflict resolution and eventual target selection are. Support for the notion that similarity between representations may tax the attentional system comes from various cognitive domains. For example, in visual categorization, increasing the similarity between the representations that need to be categorized into two sets also increases the amount of attentional control involved in the task (e.g., Sigala & Logothetis, 2002). In the language production domain, we can also find several situations in which the amount of competition between potential responses depends on their similarity. Perhaps the clearest example is the semantic interference effect in the picture–word interference paradigm. In this Stroop-like paradigm, participants have to name a target picture while ignoring the presentation of a distractor word. Naming latencies are slower when the distractor word and the target picture are categorically related than when they are unrelated, in the same way in which naming the color of the ink in which a word is written is slower if such a word corresponds to a color word than if it corresponds to an unrelated word (e.g., Caramazza & Costa, 2000; Costa & Caramazza, 2002; Glaser & Düngelhoff, 1984; Glaser & Glaser, 1989; La Heij, 1988; Lupker, 1979; Meyer, 1996; Roelofs, 1992, 1993; Rosinski, 1977; Schriefers, Meyer, & Levelt, 1990; Starreveld & La Heij, 1995, 1996; for a review of Stroop and Stroop-like paradigms, see MacLeod, 1991).

In light of these results, it is possible that the attentional system responsible for keeping the two languages apart may be more taxed when the two languages are very similar than when they are dissimilar. As a consequence, and to solve this complex situation, bilinguals of two similar languages may have developed a different attentional control mechanism than that of bilinguals with two very different languages.

Second, language similarity affects the sociolinguistic use of two languages. This is because the more similar the two languages are, the more likely people will understand both, even if they cannot speak one of them (e.g., passive bilinguals who do not speak but do understand a second language). In such a scenario, bilinguals may use their two languages at will, in almost any communicative occasion, thereby enhancing the chances for bilingual conversations. The highly proficient Spanish–Catalan bilinguals tested in our previous study are placed in precisely this sociolinguistic context. Spanish and Catalan are understood by at least 97% (and spoken by 85%) of the population of Catalonia (Viola & Moreno, 2004). Therefore, a speaker can reasonably use both languages in any conversation. In a sense, these bilinguals are “language-switching masters” as they are continuously exposed to their two languages in the same contexts. This situation is rather different when the two languages are very different because, in general, the chances for using the two languages in the same conversation are reduced.

In Experiment 1, we explore whether the language-switching performance of highly proficient bilinguals is affected by the age of L2 acquisition and language similarity. In the remaining experiments, we test the hypothesis that a difference in the proficiency level between the two languages involved in the switching task should lead to asymmetrical switching costs in highly proficient bilinguals. This hypothesis is tested under three different conditions.

In Experiment 2, Spanish–Catalan highly proficient bilinguals and learners of English perform a language-switching task involving their strong L2 and their much weaker L3. This is the first time, to our knowledge, that a language-switching task involves two “second languages.” According to the inhibitory model, asymmetrical switching costs should be expected.

In Experiment 3, Spanish–Catalan highly proficient bilinguals perform a language-switching task that involves two weak languages (participants’ L3, English, and their even weaker L4, French). The inhibitory model predicts the presence of asymmetrical switching costs because the level of proficiency is higher in L3 than in L4. Furthermore, this experiment also offers the possibility of experimentally exploring the so-called talk-foreign effect: When speaking in a weak language, the most powerful competitor seems to be another weak language rather than the strong one (De Angelis & Selinker, 2001; Ringbom, 2001; Williams & Hammarberg, 1998). (See discussion of Experiment 2.) Finally, in Experiment 4, we assess the performance of highly proficient bilinguals when the switching task involves L1 and recently learned vocabulary.

The Language-Switching Task

Before we present the experiments, it is worth discussing some general properties of the language-switching task used here. Given that some of the arguments developed in this article are based on the comparison of our results with those of other studies (Costa & Santesteban, 2004b; Meuter & Allport, 1999), it is fundamental that we follow as closely as possible the method and procedure already used in those studies. In these language-switching studies, the same reduced number of items is presented repeatedly for naming in L1 and L2 and on switch and nonswitch trials. That is, the same pictures have to be named in all four conditions. This type of design is borrowed from the task-switching studies in which switching costs in tasks of different difficulties have been explored. It is interesting that these studies have also revealed that switching into an easy task is harder than switching into a more difficult task (Allport, Styles, & Hsieh, 1994; Allport & Wylie, 1999, 2000; Yeung & Monsell, 2003; for a discussion of the conditions in which asymmetrical switching costs are observed, see Monsell, Yeung, & Azuma, 2000; and for a review of the task-switching paradigm, Monsell, 2003). Thus, this method and
procedure have proved useful to study specific issues about executive control in various cognitive domains.

However, the particularities of this design have also some drawbacks and limitations. For example, the large number of repetitions (and the fact that the same stimulus is linked to two different responses [in L1 and L2]) may create episodic effects (transient stimulus–response associations) that might affect the pattern of results. The question, however, is whether these properties interact with the presence/absence of asymmetrical switching cost. In this respect, the specific features of the present experiment seem to be less problematic. First, the number of repetitions of the stimuli does not seem to be the critical factor for observing symmetrical or asymmetrical switching costs. For example, Meuter and Allport (1999) had in their study roughly twice as many trials as Costa and Santesteban (2004b), and still the same pattern of results for L2 learners was observed in both studies. Also, in Experiment 3 of Costa and Santesteban (2004b), the number of experimental items was four times larger (40), and consequently the number of repetitions four times smaller than in their Experiment 2, and nevertheless the same pattern of results was observed. Second, when comparing the results obtained in the first half and in the second half of the experiments, a very similar pattern of language-switching costs was observed (Costa & Santesteban, 2004b). Finally, in all experiments reported below, we obtained the same pattern of results for the first 50 trials of each experiment as for the last 50 trials. Thus, although the use of such a small number of responses and their multiple repetitions may add some undesired effects, it appears that the fundamental differences between the pattern of results of L2 learners and highly proficient bilinguals are not affected by them.

Another problem with this task is that the target language is given by the stimulus rather than chosen by the speaker, a situation that short-circuits the processes involved in bilingual language production. However, this is a fundamental requirement to have experimental control over the participant’s behavior and to minimize the classical problem of exuberant responding in speech production (Bock, 1996): How can the experimenter stop participants from saying what they want? In our case, if we do not cue the language of the response, it will not be possible to have control over the language in which each trial is produced.

Therefore, and despite the limitations of this method, we think it is useful to test current models of bilingual production. In fact, the simple observation that this task is sensitive to the differential behavior of different types of bilinguals suggests that it is informing us about the nature of bilingual language control. Certainly, future research is needed to evaluate how these various parameters (e.g., different sets of stimuli for L1 and L2, larger number of stimuli per set, etc.) affect the switching performance.

Experiment 1: Language Similarity and L2 Age of Acquisition Effects in Language Switching

This experiment explores whether language similarity and age of L2 acquisition affect the language-switching performance of highly proficient bilinguals. To do so, we compare the performance of the highly proficient Spanish–Catalan bilinguals tested in Costa and Santesteban’s (2004b) study with that of (a) highly proficient bilinguals of two very dissimilar languages, Spanish and Basque (a description of the major differences between Basque and Spanish, and also about the sociolinguistic context in which these two languages are spoken, is given in Appendix A) and (b) highly proficient bilinguals who acquired their L2 (English) later in life.

Method

Participants. In Group 1, 11 native speakers of Spanish and highly proficient speakers of Basque took part in the experiment. All of the participants were undergraduate students at the University of the Basque Country (Vitoria-Gasteiz, Spain), learned Basque at school at a mean age of 3, had received education in Basque but also some courses in Spanish, and claimed to be very fluent in the two languages. They were living in the Basque Country at the time of testing and were therefore exposed to both languages. With regard to the relevant variables (education, age of L2 acquisition, years of L2 use), this sample was fully comparable with that of the Spanish–Catalan bilinguals tested in our previous work (Costa & Santesteban, 2004b, Experiment 2; see Appendix B).

Twelve Spanish–English highly proficient bilinguals participated in Group 2. They had grown in a Spanish monolingual context (outside of Catalonia) and started acquiring the L2 relatively late in life. Furthermore, they learned English in the classroom and did not start using English more or less regularly until the age of 15 (for more details, see Appendix B). Nevertheless, they attained a high proficiency level in English, because they kept studying it at university and were enrolled in translation and interpretation programs, becoming professional interpreters and translators.2 All of the participants had received university education and were living in Barcelona at the time of testing.

Materials. The materials, design, and procedure were almost identical to those used by Costa and Santesteban (2004b). For Group 1, 10 pictures of common objects with noncognate names were selected. For Group 2, we used the same materials as in Costa and Santesteban’s (2004b) Experiment 4 (see Appendix C for a description of the whole set of materials). Participants were instructed to name the picture in Spanish or Basque (or Spanish or English for Group 2) based on the color of the picture (red or blue). The assignment of color cue to response language was counterbalanced across participants. Pictures were presented in short sequences (“lists”) ranging in length, unpredictably, from 5 to 14 trials. There were two types of trials: (a) trials on which the language of the response (either L1 or L2) was the same as on the trial immediately before (nonswitch trials) and (b) trials on which the language of the response (either L1 or L2) was different from the one used on the preceding trial (switch trials). The lists varied in the number of switch trials (from 0 to 4 switch trials). A total of 400 such lists was constructed. These lists varied in the specific sequence of (a) L1 and L2 responses and (b) switch and nonswitch trials. Each participant completed 100 lists for a total of 950 trials (70% non-switch and 30% switch). On half of the nonswitch trials, participants were asked to name the picture in L1 (333 trials) and on the other half, in L2 (333 trials). The same applied to the switch trials. Therefore, participants used their L1 and L2 an equal amount of time throughout the course of the experiment (475 responses in each language). For each participant, 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

2 The reason for choosing this particular population is the large negative correlation between the age at which L2 is acquired and language proficiency. That is, in general, the earlier a speaker learns an L2, the higher his or her proficiency level is. So, to ensure as much as possible that the late learners were as proficient as the early learners in their L2, we followed the same strategy already used in other bilingual studies in which the contribution of proficiency and age of acquisition to language processing has been explored (Perani et al., 1998; but see Christoffels & De Groot, 2004). Thus, arguably, we chose those late bilinguals with the highest proficiency level that one could find.
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 on switch trials than on nonswitch trials for a 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 2 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 and as accurately as possible. 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. Before the experiment began, participants were familiarized with the pictures’ names in the two languages. In the familiarization phase the experimental pictures were presented in black outline, and participants were asked to name them in L1, and afterward in L2. This familiarization was repeated three times. A list of experimental trials had the following structure: (a) a blue or red circle along with the word Español (Spanish) or Euskera (Basque)— or Inglés (English) for Group 2—written below was presented for 2,000 ms (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 nonswitch trial); (b) the first picture of a list appeared and remained on the screen for 2,000 ms or until participants responded; (c) a blank interval of 1,150 ms; (d) the next picture was presented, and the cycle was repeated until the end of the list; and (e) after the presentation of the last picture of the list, an asterisk was presented for 1,000 ms signaling the end of the list. The experiment started with the presentation of six training lists, which were excluded from subsequent analyses.

Results

Error rates and naming latencies were submitted to two analyses of variance with two variables: response language (L1 vs. L2) and type of trial (switch vs. nonswitch). Although analyses by items are not usually reported in this type of experiment (given the small number of items and their random assignment to each condition), for the sake of completeness we present them here. Two types of responses were scored as errors: (a) verbal disfluencies (stuttering, utterance repairs, etc.) and (b) voice key failures. Also, naming latencies exceeding 3 standard deviations 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. The results of Group 1 and 2 were analyzed separately.

Group 1: Spanish–Basque bilinguals. Error analyses were performed over all of the excluded trials (5.7% in total; errors: 4.1%; outliers: 1.6%). Averages of the different conditions along with the associated 95% confidence interval can be found on Figure 1.

Participants made more errors in their L1 (6.9%) than in their L2 (5%), $F(1, 10) = 5.78, MSE = 6.99, p = .037$; $F(2, 9) = 6.78, MSE = 4.96, p = .028$; min $F(1, 19) = 3.12, p = .093$; and on the switch trials than on the nonswitch trials (6.6% and 5.4%, respectively), $F(1, 10) = 5.96, MSE = 2.14, p = .035$; $F(2, 9) = 2.6, MSE = 4.99, p = .141$; min $F(1, 16) = 1.81, p = .197$. The interaction of these two variables was not significant (both $F$s < 1).

Faster naming latencies were observed for L2 (703 ms) than for L1 trials (776 ms), $F(1, 10) = 30.32, MSE = 1.943.01, p < .001$; $F(2, 9) = 20.57, MSE = 2.665.02, p = .001$; min $F(1, 18) = 12.25, p = .002$. Switch trials (752 ms) led to longer naming latencies than nonswitch trials (727 ms), $F(1, 10) = 27.94, MSE = 253.2, p < .001$; $F(2, 9) = 33.3, MSE = 222.95, p < .001$; min $F(1, 19) = 15.19, p = .001$. More important, the interaction between the two variables was not significant (both $Fs < 1$), revealing that the magnitude of the switching cost was comparable for the two languages (L1: 21 ms; L2: 29 ms).

The results of Group 1 were compared with those of the Spanish–Catalan bilinguals tested in Costa and Santesteban (2004b, Experiment 2; see Figure 2). The overall pattern of results was similar for both groups as revealed by the lack of interaction between response language (L1 vs. L2), type of trial (switch vs. nonswitch), and group of participants (Spanish–Basque vs. Spanish–Catalan), all $Fs < 1$.

Group 2: Spanish–English late bilinguals. Following the criteria presented above, 6.1% of the trials (errors: 4.5%; outliers: 1.6%)
were excluded from the analyses. The error analysis showed significant effects of response language (L1: 6.9%; L2: 5.2%), \( F(1, 11) = 8.29, \text{MSE} = 5.81, p = .015; F(2, 1, 9) = 5.49, \text{MSE} = 7.53, p = .044; \) min \( F'(1, 18) = 3.30, p = .085; \) and type of trial (switch: 7.3%; nonswitch: 5.5%), \( F(1, 11) = 17.74, \text{MSE} = 2.22, p = .001; F(2, 1, 9) = 11.56, \text{MSE} = 2.79, p = .008; \) min \( F'(1, 18) = 6.99, p = .01. \) The interaction of these two variables was not significant.

In the analysis of naming latencies, the main effects of response language (L2: 693 ms; L1: 720 ms), \( F(1, 11) = 7.14, \text{MSE} = 1,198.33, p = .022; F(2, 1, 9) = 9.38, \text{MSE} = 832.35, p = .013; \) min \( F'(1, 20) = 4.05, p = .05; \) and type of trial (nonswitch: 686 ms; switch: 727), \( F(1, 11) = 44.71, \text{MSE} = 432.05, p = .001; F(2, 1, 9) = 207.81, \text{MSE} = 77.95, p = .001; \) min \( F'(1, 15) = 36.79, p = .001, \) were significant (see Figure 3). More important, the interaction between the two variables was not significant, \( F(1, 9) < 2.43, \text{MSE} = 92.84, p = .153, \) revealing that the magnitude of the switching cost was comparable for the two languages (L1: 43 ms; L2: 37 ms).

When considering together the results of the Spanish–English late bilinguals with those observed with Spanish–Catalan early bilinguals, the crucial interaction between response language (L1 vs. L2), type of trial (switch vs. nonswitch), and group of participants (Spanish–English vs. Spanish–Catalan) was not significant (all \( Fs < 1 \)), revealing that the pattern of switching costs was similar for both populations.

Discussion

The goal of this experiment was to assess the impact of language similarity and L2 age of acquisition in the language-switching performance of highly proficient bilinguals. Overall, the results from Spanish–Basque highly proficient early bilinguals and Spanish–English highly proficient late bilinguals resemble very much those observed for the Spanish–Catalan highly proficient early bilinguals. The following results were found for these bilinguals: (a) The magnitude of the switching cost was similar for the two languages, and (b) naming latencies were faster for L2 than for L1.\(^3\) Thus, it appears that neither language similarity nor age of L2 acquisition affects in a significant manner the way bilinguals control their speech, at least as indexed by the language-switching task.

These results are consistent with the inhibitory control model (Green, 1998), according to which asymmetrical switching costs should emerge only when the difference in proficiency between the two languages involved in the switching task is large. However, as advanced in the introduction (see Footnote 1), Costa and Santesteban (2004b) reported one result that seems at variance with this prediction: Spanish–Catalan highly proficient bilinguals performing the switching task in their L1 and in their much weaker L3 (English) did not show asymmetrical switching costs. In fact, their performance when switching between two dominant languages (Spanish–Catalan) or between a dominant language (L1: Spanish) and a much weaker one (L3: English) was very similar. This observation is problematic for the inhibitory model because, if we were to interpret the asymmetrical switching costs observed for L2 learners as revealing the differential levels of inhibition applied to their two languages, then we should observe such an effect for highly proficient bilinguals performing the task in L1 and in the much weaker L3.

Given the important implications of this observation for the interpretation of the language-switching performance of bilinguals, we aim at replicating and extending it in Experiment 2. In this experiment, Spanish–Catalan highly proficient bilinguals and learners of English were asked to perform the switching task in their strong second language (L2) and their much weaker third language (L3). Furthermore, this experiment also serves as a first assessment of the language-switching performance when two second languages are involved in the task, an issue that is explored in more detail in Experiment 3.

\(^3\) We defer further discussion of this language effect to the General Discussion.
Experiment 2: Highly Proficient Bilinguals Switching Between L2 and L3

In this experiment we aim at replicating and extending the observation of symmetrical switching costs when highly proficient bilinguals are asked to switch between a dominant and a weak language. We do so by asking Spanish–Catalan bilinguals to switch between their L2 (Catalan) and their much weaker L3 (English). This is an interesting and novel condition, given that up to now all language-switching experiments have involved bilinguals’ first language. Thus, this experiment will also provide us with relevant information about the role played by the presence of the L1 in the language-switching task.

If a difference in proficiency between the two languages is at the basis of the asymmetrical switching cost, then we should find such a pattern in this experiment. Alternatively, if highly proficient bilinguals do not make use of the same inhibitory mechanism as used by L2 learners (as suggested by Costa & Santesteban, 2004b), then we should find symmetrical switching costs, as observed in Experiment 1.

Method

Twelve highly proficient Spanish–Catalan bilinguals took part in the experiment (see Appendix B). All of the participants had taken English courses for at least 8 years (1-hr classes twice a week) as part of their high school curriculum. However, only 1 of them reported having subsequently enrolled in an English course, only 2 reported having been to an English-speaking country (for a period not longer than 2 weeks), and all except 4 claimed they did not use English at all at the time of testing. As the language proficiency self-assessment shows, these participants’ L3 proficiency levels are even lower than those of the L2 learners tested by Costa and Santesteban (2004b, Experiment 1). Consequently, participants in this group can be considered L3 learners (people with some L3 experience but not very high proficiency). The materials, design, and procedure were identical to that of Experiment 2. The name of the 10 pictures presented had noncognate names (see Appendix C for a description).

Results and Discussion

Following the same criteria as in Experiment 1, 5.6% (errors: 4.2%; outliers: 1.4%) of the trials were excluded from the analyses. The only significant effect in the error analysis was that of type of trial (switch: 7.4%; nonswitch: 4.9%), $F(1, 11) = 24.79, MSE = 2.99, p < .001; F(2, 1, 9) = 21.31, MSE = 3.12, p = .001; min F'(1, 1) = 11.45, p = .004$.

In the analysis of naming latencies, the main effect of response language was significant in the participants’ analysis (L3: 871 ms; L2: 913 ms), $F(1, 11) = 7.34, MSE = 2,849.02, p = .02; F(2, 1, 9) = 3.02, MSE = 5,178.35, p = .16; min F'(1, 16) = 2.13, p = .16$. The main effect of type of trial was also significant (switch: 864 ms; nonswitch: 919 ms), $F(1, 11) = 59.33, MSE = 613.62, p < .001; F(2, 1, 9) = 29.1, MSE = 1,052.55, p < .001; min F'(1, 17) = 19.52, p = .001$. The interaction between the two variables was not significant (both Fs < 1; switch to L2: 57 ms; switch to L3: 53 ms; see Figure 4).

The results of this experiment are clear: When highly proficient bilinguals are asked to switch between their strong L2 and their much weaker L3, symmetrical switching costs are observed. This result, together with that of Costa and Santesteban’s (2004b) Experiment 4, reveals that differences in the proficiency levels between the two languages involved in the switching task do not lead to asymmetrical switching costs for highly proficient bilinguals. This suggests that, when performing the language-switching task, they do not use the same control mechanism (at least as indexed by such a task) as L2 learners.

These results are consistent with the notion that an increase in L2 proficiency may lead to a shift in the selection mechanisms that guarantee lexical selection in the intended language. Costa and Santesteban (2004b) stated this hypothesis as follows:

[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 . . . .] 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 . . . .
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. (p. 505)

Accordingly, highly proficient bilinguals should not show asymmetrical switching costs, regardless of the strength of the two languages involved in the switching task.

In the following two experiments, we test the functionality of this hypothesized language-specific selection mechanism in other experimental contexts. In Experiment 4, highly proficient Catalan–Spanish bilinguals performed a switching task involving two languages (their L3 is English and L4 is French) for which they had lower proficiency levels than for their two dominant languages. This is an interesting condition not only because this is the first time a language-switching task involves two weak languages but also because of the following reason. Research with multilingual speakers has demonstrated that trilinguals with low L3 proficiency often show more interference from their weak L2 than from their dominant L1 (De Angelis & Selinker, 2001; Ringbom, 2001; Williams & Hammarberg, 1998). It is as if they can control very efficiently the potential interference from their dominant language(s), but they have more difficulties in doing so with the interference produced by a weak language. If this observation actually reveals the use of a different control mechanism, perhaps highly proficient bilinguals performing the task in two weak languages (their L3 and L4) cannot make use of the language-specific selection mechanism and need to resort to other mechanisms. In other words, it is possible that the language-specific selection mechanism is only functional for highly proficient bilinguals when one of the languages involved in the switching task is a strong one (either their L1 or their L2).

Experiment 3: Exploring Further the Language-Switching Performance of Highly Proficient Bilinguals: Switching Between L3 and L4

This experiment aims at exploring the language-switching performance of highly proficient bilinguals when switching between two weak languages. According to the hypothesis that highly proficient bilinguals should not be affected by the difference in the proficiency levels of the two languages involved in the switching task, symmetrical switching costs are predicted.

### Method

Twelve Spanish–Catalan highly proficient speakers took part in the experiment (see Appendix B). All of the participants had studied English for a minimum of 8 years as part of their high school curriculum at the time of testing and French for at least 1 year \((M = 3.14, SD = 1.8)\), with the exception of 1 participant who reported having studied it for 70 hr, without specifying the time period over which this took place. All of the participants reported to be much less proficient in these two languages than in Spanish and Catalan.

A set of 10 pictures with noncognate names in all four languages was selected (see Appendix D). The design and procedure were the same as in Experiment 1.

### Results and Discussion

Following the same criteria presented above, 6.3% (errors: 5.1%; outliers: 1.2%) of the trials were excluded from the analyses. In the error analysis, the main effect of type of trial approached significance (switch: 7.2%; nonswitch: 6%), \(F(1, 11) = 3.66, MSE = 4.45, p = .082\); \(F(2,1, 9) = 11.11, MSE = 1.06, p = .009\); \(min F'(1, 17) = 2.75, p = .11\); and that of response language was significant (L4: 8.7%, L3: 4.4%), \(F(1, 11) = 22.03, MSE = 8.85, p = .001\); \(F(2,1, 9) = 14.16, MSE = 12.2, p = .004\); \(min F'(1, 15) = 8.61, p = .009\).

In the analysis of naming latencies, the main effects of type of trial (nonswitch: 847 ms; switch: 963 ms), \(F(1, 11) = 28.96, MSE = 690.69, p < .001\); \(F(2,1, 9) = 52.19, MSE = 326.78, p < .001\); \(min F'(1, 19) = 18.62, p = .001\); and response language were significant (L3: 885 ms, L4: 926 ms), \(F(1, 11) = 38.78, MSE = 4,145.65, p < .001\); \(F(2,1, 9) = 24.34, MSE = 5,450.62, p = .001\); \(min F'(1, 18) = 14.95, p = .001\). The interaction between the two variables was also significant, \(F(1, 11) = 27.35,
The results of this experiment are clear: The magnitude of the switching cost was four times larger for L3 than for L4. That is, when highly proficient bilinguals perform the switching task in two languages for which their proficiency is relatively low, the magnitude of the switching cost is larger for the stronger language of the two.

This observation is at odds with the prediction derived from Costa and Santesteban’s (2004b) hypothesis, according to which highly proficient bilinguals should not show asymmetrical switching costs under any circumstances. These results might be revealing certain limits in the cognitive flexibility of the bilinguals’ switching mechanism, suggesting that the language-specific selection mechanism is not available when the switching task does not involve one of the strong languages of these bilinguals. Before drawing any strong conclusion, we further test Costa and Santesteban’s hypothesis by asking bilinguals to switch between their L1 and a recently learned set of words belonging to a new language.

Experiment 4: Controlling Speech Production in the First Stages of Word Learning

In this experiment, we explore the language-switching performance of two groups of participants: Spanish monolinguals and Spanish–Catalan highly proficient bilingual speakers. Before the switching task, participants were asked to learn 10 novel words, coined by the experimenters and supposedly belonging to a “new language.” After the learning phase (see Procedure below), participants performed a language-switching task between their L1 and the new language.

If, as hypothesized by Costa and Santesteban (2004b), highly proficient bilinguals have developed a control mechanism that does not resort to inhibition and such a mechanism is functional no matter the level of proficiency of the two languages involved in the switching task, then this type of bilingual should show symmetrical switching costs while L2 learners should show asymmetrical ones.

Method

Participants. Participants in Group 1 (n = 12) were Spanish monolingual university students with little knowledge of another language. Although they had formally learned English or French at school as a foreign language, none of them reported using L2 on a daily basis or having high proficiency in that language. This group could in fact be considered L2 learners, but for the sake of clarity, we refer to them as Spanish monolinguals. Participants in Group 2 (n = 12) were recruited from the same population of that reported in Experiment 2 of Costa and Santesteban (2004b). All of them were highly proficient Spanish–Catalan early bilinguals and were students at the University of Barcelona (see Appendix B). Participants in both groups were comparable in age and education.

Procedure. The same method, materials, and procedure as in Experiment 1 were used. However, before the switching task, participants were required to learn 10 invented words corresponding to each of the 10 pictures (for more details, see Appendix C). The learning phase had three parts. First, all pictures were presented five times with the corresponding name in the new language written below and played through headphones, and participants were asked to listen and repeat them. Second, the pictures were presented three times with their names written below, and participants were asked to produce the name of each picture. Third, pictures were presented separately, and participants were asked to name them.

Results and Discussion

Following the criteria presented above, 5.2% and 5.7% of the trials for the Spanish monolingual group (errors: 3.5%; outliers: 1.7%) and for the Spanish–Catalan bilingual group (errors: 4.4%; outliers: 1.3%) were removed from the analyses. The main effect of response language (new language: 8.8%; L1: 2.1%), $F_{1(22)} = 102.83$, $MSE = 9.68$, $p = .001$; $F_{2(1, 9)} = 41.3$, $MSE = 20.34$, $p = .001$; $min F_{(1, 17)} = 29.46$, $p = .001$, was significant. All other main effects were not significant. The interaction be-
between response language and type of trial was significant, \(F(1, 22) = 7.12, \text{MSE} = 1.6, p = .014\); \(F(1, 9) = 3.88, \text{MSE} = 2.11, p = .08\); \(\text{min } F(1, 19) = 2.51, p = .12\), revealing that the difference in error rates between switch and nonswitch trials was larger in L1 (2.8% and 1.8%) than in the new language (8.5% and 8.9%, respectively).

In the analyses of naming latencies, the main effects of response language, \(F(1, 22) = 78.41, \text{MSE} = 9.306.54, p = .001\); \(F(1, 9) = 74.27, \text{MSE} = 8.059.47, p = .001\) and \(\text{min } F(1, 25) = 38.14, p = .001\); and type of trial, \(F(1, 22) = 70.83, \text{MSE} = 591.72, p = .001\); \(F(1, 9) = 132.05, \text{MSE} = 286.58, p = .001\); and \(\text{min } F(1, 30) = 46.10, p = .003\), were significant. These results showed that participants were faster naming pictures in L1 than in the new language (755 and 929 ms, respectively), and faster on nonswitch (821 ms) than on switch trials (863 ms). The main effect of group of participants was only marginally significant in the analysis by participants, \(F(1, 22) = 3.04, \text{MSE} = 35.940.13, p = .095\); \(F(1, 9) = 70.93, \text{MSE} = 1,231.83, p < .001\); \(\text{min } F(1, 24) = 2.91, p = .10\). Importantly, the interaction between response language and type of trial was significant, \(F(1, 22) = 41.37, \text{MSE} = 290.4, p = .001\); \(F(1, 9) = 22.93, \text{MSE} = 399.33, p = .001\); \(\text{min } F(1, 19) = 14.75, p = .001\). No other significant interactions were observed (see Figures 6 and 7).

Furthermore, when the results of the two groups of participants were analyzed separately, the crucial interaction between response language and type of trial was significant for both: monolinguals, \(F(1, 11) = 20.77, \text{MSE} = 396.83, p = .001\); \(F(1, 9) = 19.76, \text{MSE} = 344.71, p = .002\); \(\text{min } F(1, 20) = 10.12, p = .005\); and bilinguals, \(F(1, 11) = 22.42, \text{MSE} = 183.97, p = .001\); \(F(1, 9) = 11.56, \text{MSE} = 241.17, p = .008\); \(\text{min } F(1, 17) = 7.67, p = .013\). This interaction shows that the magnitude of the switching cost was larger for L1 than for the new language (L1: 65 and 63 ms; new language: 13 and 26 ms, for monolinguals and bilinguals, respectively).

The results of this experiment reveal that highly proficient bilingual and monolingual speakers show asymmetrical switching costs when the switching task involves their L1 and words from a recently learned new language. That is, the magnitude of the switching cost was 2.5 times larger for the dominant language (L1) than for the new language. This is the first time that highly proficient bilinguals and monolinguals show the same performance in a language-switching task, suggesting that the mechanisms used by both populations for controlling speech production at very early stages of word learning are the same.

**General Discussion**

We reported four experiments exploring the language-switching performance of highly proficient bilinguals in various contexts. In Experiment 1, the language-switching performance of two groups of bilinguals was assessed: (a) highly proficient bilinguals of two very dissimilar languages (Spanish–Basque) and (b) Spanish–English highly proficient bilinguals who learned their L2 late in life. The results of both groups were very similar to those of highly proficient early bilinguals of two very similar languages, Catalan–Spanish (Costa & Santesteban, 2004b): Language-switching costs were of comparable magnitude for the bilinguals’ two languages (symmetrical switching costs).

These results revealed that the language-switching performance of highly proficient bilinguals is not affected by either language similarity or the age at which L2 was acquired. Thus, we can safely conclude that when the language-switching task is performed in two languages in which the bilingual is very proficient, symmetrical switching costs are observed. This pattern contrasts sharply with that shown by L2 learners (see Figure 6), for whom asymmetrical switching costs are observed, also despite language similarity.

In the remaining three experiments, we further tested some predictions derived from the inhibitory model, and the results revealed a complex scenario. In Experiment 2, symmetrical switching costs were observed when highly proficient bilinguals were asked to switch between their strong L2 and their much weaker L3. These results are clearly at odds with the inhibitory model and replicate and extend our previous observations (Costa...
suggesting that symmetrical switching will always be present for those bilinguals who have acquired a high-proficiency level in any pair of languages. In fact, such an observation indicates that highly proficient bilinguals do not resort to inhibition during lexical access in speech production. Instead, we claimed that they make use of a language-specific lexical selection mechanism that considers for selection only the lexical representations corresponding to the intended language. We further argued that once this language-specific selection mechanism kicks in, bilinguals can make use of it regardless of the strength of the languages involved in the language-switching task.

However, the results of Experiments 3 and 4 showed this latter conclusion to be wrong. In Experiment 3, highly proficient bilinguals were asked to switch between two languages in which their proficiency was lower than that of their dominant languages (English: L3 and French: L4). For the first time, highly proficient bilinguals showed asymmetrical switching costs. In Experiment 4, monolinguals and bilinguals performed a language-switching task involving L1 and a recently learned vocabulary. Both groups showed asymmetrical switching costs (larger switching cost for their L1 than for the new language). Thus, these two experiments reveal the presence of asymmetrical switching costs for highly proficient bilinguals, rejecting therefore Costa and Santesteban’s (2004b) hypothesis.

Given the whole set of results, and assuming that asymmetrical switching costs index the presence of inhibitory control, we should conclude that in some conditions highly proficient bilinguals resort to this mechanism and in others they do not. The critical issue now is whether this complex pattern of results allows us to draw any theoretically relevant interpretation about the nature of the lexical selection mechanism in highly proficient bilinguals. For one thing, the available experimental evidence reveals a more complex scenario than that considered by Costa and Santesteban (2004b) when they hypothesized a shift in the lexicalization processes linked to an increase in L2 proficiency. And, given this complexity, one could be tempted to avoid drawing any other interpretation from these results. However, if we are able to identify which properties of the language-switching task are behind the presence or absence of asymmetrical switching costs, we may find a reasonable, albeit tentative, interpretation of these results.

**Determining the Source of Asymmetrical Switching Costs in Highly Proficient Bilinguals**

Let us first point to some properties of the switching task that are irrelevant for the presence of asymmetrical switching costs in highly proficient bilinguals (see Figure 8): (a) the differential level of proficiency between the two languages, (b) the similarities between the languages, (c) the age at which the second language has been acquired, and (d) whether the L1 or the L2 is involved in the switching task. This is because symmetrical switching costs are observed when the difference in the proficiency levels is large and when it is small (Experiments 1 and 2), when the switching task involves two very similar languages (Spanish–Catalan) and when it involves two very dissimilar languages (Spanish–Basque; Experiment 2 in Costa & Santesteban, 2004b, and current Experiment 1, Group 1), when the second language has been acquired early in life but also when it has been acquired later (Experiment 2 in Costa & Santesteban, 2004b, and current Experiment 1, Group 2), and when L1 is present in the task but also when L2 rather than L1 is present (Experiment 4 in Costa & Santesteban, 2004b, and current Experiment 2).

Let us focus now on the common feature(s) of those experimental conditions leading to asymmetrical switching costs in highly proficient bilinguals (Experiments 3 and 4). These experiments always involved a rather weak language. That is, the lexical representations of one of the two languages involved in the switching task were not very well established (either L4 or the recently learned vocabulary from the new language). Thus, one could make the following empirical generalization: Whenever the language-switching task involves a language for which the corresponding words are not well established, highly proficient bilinguals will show asymmetrical switching costs, otherwise symmetrical switching costs will be obtained. However, this empirical generalization is only interesting from the theoretical point of view if we
can find independent reasons to argue that such a variable can affect the control mechanism used by bilinguals to perform the language-switching task. There are two reasons why this could be the case, and both are related to whether the language-specific selection mechanism can operate on the lexical representations.

First, the robustness (familiarity and frequency leading to greater automaticity of retrieval) of the lexical representations of the weakest language may be crucial for the functionality of the language-specific selection mechanism. Such a mechanism is assumed to guarantee lexical selection in the intended language by inspecting the activation of only those representations belonging to the language that has been chosen for production (the response language). If the lexical representations of the response language are integrated in a lexicon (i.e., represented in the long-term memory in an integrated lexicon; see Kroll & de Groot, 1997; Talamas, Kroll, & Dufour, 1999; Van Hell & De Groot, 1998, for the variables that can affect such an integration), then the lexical selection mechanism can focus on that lexicon and avoid interference from the other one, without the need of inhibiting the lexical representations of the nonresponse language. That is, the language-specific selection mechanism can only be sensitive to those representations that are integrated into a lexicon. When this criterion is not met, as is presumably the case at very early stages of word learning or with an L4, it is impossible for the selection mechanism to perform a restricted search over sparse lexical representations that do not belong to a specific lexicon. In short, the language-specific selection mechanism can only be functional when there actually is a language-specific lexicon to focus on.

However, participants are able to perform a switching task involving a rather weak language. How do they do that? How do they manage to prevent lexical intrusions from their strong language(s) when speaking in the weak language? One possibility is that they prevent interference by actively inhibiting the lexical representations of their strong language(s). Note that the inhibitory mechanism can function with no problem when a very weak language is involved in the switching task, because it operates primarily over the representations of the strong language. That is, this mechanism ensures the correct selection in the weak language by inhibiting the lexical representations belonging to the language that has not been chosen for production (the nonresponse language). Therefore, this mechanism can be fully functional when producing the recently learned words of the new language (or L4), given that it operates on the well-established lexical representations.4

There is another reason of why the robustness of the lexical representations of the weakest language may be crucial for the functionality of the language-specific selection mechanism. It has

---

4 For this argument to work, we should assume that the lexical representations of the L3 (English) of the bilinguals tested in Experiment 4 are already integrated into a lexicon that allows the inhibitory mechanism to function on them. This assumption receives some support from the results of Experiment 3. In that experiment in which Spanish–Catalan bilinguals were asked to switch between their L2 and their L3 (English), we observed symmetrical switching costs. If we were to interpret the presence of symmetrical switching costs as revealing that the lexical representations of the L3 are robust enough to allow the language-specific selection mechanism to kick in, then we should conclude that they are also robust enough for the inhibitory mechanism to work as well. And given that the L3 of the participants in Experiment 3 was of similar strength to the L3 of the participants in Experiment 4, then we could assume that in both cases the lexical representations of the L3 were established enough for the inhibitory mechanism to work on them.
long been argued (e.g., Kroll & Stewart, 1994) that when the L2 of a bilingual is not well established yet, access to its lexical representations is mediated by previous access to the corresponding translation from the strong language. Once the bilingual increases her or his proficiency in the weak language, the lexical representations of that language establish direct links with the conceptual system, allowing for a direct access to these lexical representations without need of L1 mediation. What would be the consequences of such an architecture for the language-switching task? If L2 production involves L1 mediation, the language-specific selection mechanism cannot operate. This is because the lexical representations of the nonresponse language (L1) need to be selected to produce those in the response language (L2). Thus, a lexical selection mechanism that ignores the activation of the L1 representations when producing words in the weak language does not seem viable. However, in such circumstances, the inhibitory control mechanism may still be functional. This is because, at some point during the lexicalization process in the weak language, the bilingual needs to avoid interference from the lexical representation of the strong language. Once such an L1 representation has been selected and translated into the weak language, it needs to be suppressed so that its overt production is prevented. In this framework, naming in the weak language would entail (a) retrieval of the lexical representations of the strong language, (b) translation into the weak language, and (c) inhibition of the lexical representations of the strong language to prevent their overt production. This last step would be behind the asymmetrical switching costs observed when the task involves one language in which the bilingual speaker has very low proficiency.

The explanations given above for the presence of asymmetrical switching costs in highly proficient bilinguals are based on whether the language-specific selection mechanism can operate over the representations of the weak language. Accordingly, the greater the robustness of these lexical representations, the more likely successful selection of those representations will be achieved without the need of inhibitory mechanisms. Thus, for highly proficient bilinguals, what matters for the presence or absence of asymmetrical switching costs is not the difference in proficiency levels between the two languages involved in the switching task but rather how well established the weak language involved in the task is. Admittedly, and although this characterization captures the available experimental evidence, it is rather tentative, and future research needs to assess its validity.

Despite the merits of this explanation of the performance of highly proficient bilinguals, there is still the question of why highly proficient bilinguals and L2 learners behave differently under similar experimental conditions. That is, highly proficient bilinguals show symmetrical switching costs (Experiment 2) under the same conditions (in terms of the difference in strength of the two languages and in terms of how well established the weak language is) in which low-proficient bilinguals show asymmetrical switching costs. As we argued in Costa and Santesteban (2004b), we believe that this contrasting pattern of results has its origin in the development of language-specific selection mechanisms by highly proficient bilinguals. In contrast, low-proficient L2 learners need to make use of inhibitory mechanisms even when the lexical representations of the weak language are well established.

Inhibitory Control: Item Specific or Language Specific?

Some remarks about the nature of the inhibitory control are pertinent here. In the arguments developed through this article, we have assumed that the inhibitory control mechanism affects the nonresponse language as a whole. That is, when naming in one language, all the lexical representations of the nonresponse language are inhibited. However, given that we have used a limited set of target stimuli, one may claim that the inhibitory mechanism acts only on some specific lexical representations rather than on the whole nonresponse language. Unfortunately, our results are not helpful to adjudicate between these alternatives. However, the results from other similar studies suggest that, at least, part of the inhibition affects the nonresponse language as a whole (Loasby, 1998; Meuter & Allport, 1999). Nevertheless, a resolution of this issue will only be obtained by manipulating the stimuli included in the language response set. That is, we need to construct conditions in which the pictures that have to be named in Language A are different from those that have to be named in Language B. We believe that such an experiment falls outside the scope of the present article (for preliminary results, see Wodniecka, Bobb, Kroll, & Green, 2005).

Symmetrical Switching Costs Are Associated With L2 Latency Advantage Over L1

It is worth discussing another result observed in our study: Naming latencies in the weak language are lower than in the strong language. This is a paradoxical but very robust observation found in all language-switching experiments involving highly proficient bilinguals (see also Jansen, 1999). At present, the only explanation we have for the L2 advantage over the L1 is the one proposed by Costa and Santesteban (2004b). We argued that when bilinguals develop a language-specific selection mechanism, they may also be able to set different selection thresholds for their two languages, allowing them, in language-switching tasks, to prioritize selection in the weakest language of the switching task. Note, however, that such a difference may also be revealing the presence of an inhibitory mechanism of a different sort to that indexed by the asymmetrical switching cost. That is, one may claim that responses in the strong language are slower because the corresponding representations have been inhibited in a general manner across the whole set of trials. At present, this inhibitory mechanism and the different selection threshold seem to be quite similar and difficult to put to test.

The results of Experiments 3 and 4 add interesting information about the conditions that have to be met for this language effect to arise. In these experiments, naming latencies were lower in the strong than in the weak language, and at the same time, asymmetrical switching costs were observed. Thus, there is an instructive association of effects: When asymmetrical switching costs are observed, latencies in the strong language are higher than in the weak language (e.g., Experiments 1 and 2), and when asymmetrical switching costs are observed, latencies in the strong language are lower than in the weak language (Experiments 3 and 4).

These associations are consistent with the explanation proposed by Costa and Santesteban (2004b) of the language effect. This is because, if the presence of symmetrical switching costs actually reveals the functionality of the language-specific selection mech-
anism, and the advantage of the weak language is due to the flexibility given by such a mechanism to set different selection criteria thresholds, then the two effects should go hand in hand. Furthermore, we should also predict that in those conditions in which the lexical-selection mechanism cannot kick in, and, as a consequence, different selection thresholds cannot set up, naming latencies in the strong language should be faster and asymmetrical switching costs should be present. At any rate, and despite the fact that this observation is consistent with the only proposed explanation of the paradoxical advantage of the weak over the strong language, further research needs to assess the precise theoretical implications of this effect. Our contribution here is the observation that in highly proficient bilinguals the advantage of the weak over the strong language disappears when asymmetrical switching costs are observed.

Conclusion

Bilingual speakers need to acquire control mechanisms to restrict their lexicalization process to one of their languages while preventing massive interference from the other language. It has been proposed that highly proficient bilinguals make use of a different control mechanism compared with low-proficient bilinguals. The first type of speaker guarantees selection by means of a language-specific selection mechanism, whereas the latter guarantees selection by means of inhibitory control. The present results show that the mechanism used by highly proficient bilinguals does not depend either on the linguistic similarity between the two languages or on the age at which the L2 has been acquired. Our results also reveal some limits on the functionality of the language-specific selection mechanism and suggest that in some conditions, highly proficient bilinguals need to resort to inhibitory control to perform lexical selection. Future research is needed to test the validity of the explanations put forward in this study to account for the complex experimental evidence.

References


LANGUAGE SWITCHING IN HIGHLY PROFICIENT BILINGUALS


Sebastián-Gallés, N., & Bosch, L. (2002). The building of phonotactic knowledge in bilinguals: The role of early exposure. *Journal of Exper-
imental Psychology: Human, Perception and Performance, 28, 974–989.

Appendix A

Main Differences Between Basque and Spanish and the Sociolinguistic Situation in the Basque Country

Basque and Spanish are very different at the lexical, morphological, and syntactic levels. Basque is a non-Indo-European language typologically very far away from Romance languages (e.g., Spanish, Catalan, Italian, French). Although Basque has been in contact with Romance languages (Latin and Spanish) in the course of many centuries and has taken and adapted words from them, it remains different in many linguistic aspects. For example, Basque is an agglutinative language that could attach to the lexical roots morphemes with syntactic information like number, case, aspect, mode, and so on. In such a way, large words are created in Basque. In contrast, Spanish and Catalan are flexive languages. These languages attach to the lexical roots derivational suffixes or gender, number, mode, or verb aspect inflexions. In addition, Basque is an ergative language, whereas Spanish and Catalan are accusative languages. Basque is a head-final parameter language that establishes the subject–object–verb order as its canonical sentence word order, whereas Spanish is a head-initial parameter language (like English) with subject–verb–object as the canonical sentence word order. Basque has postpositions, whereas Spanish has prepositions. Basque does not mark gender for common nouns whereas Spanish does (for a brief description of Basque grammar, see Laka, 1996). All these factors, among others (e.g., percentage of cognate words), make Basque much harder to learn than Catalan for a native Spanish speaker.

Regarding the sociolinguistic situation in the Basque Country, it is worth noting that the percentage of bilingual individuals in the Basque Country is relatively low. From the population of above 15 years of age, only about 25% are balanced bilinguals, whereas about 10% are passive bilinguals who do not speak Basque fluently but can understand it. Then, at least about 64% of the population does not understand Basque (Basque Government, 2003). This scenario certainly reduces the chances that a conversation between several people can be conducted switching between the two languages, as some of the interlocutors would likely be excluded because of the use of Basque. On such occasions, Spanish (which is spoken by everyone in the Spanish part of the Basque Country) or French (in the case of the part of the French Basque Country) becomes the language used in a conversation (see Etxebarria, 2002, for a detailed comparative description of the sociolinguistic situation of the Basque Country and Catalonia). In short, although L2 proficiency might be, a priori, independent of the similarity between languages, the fact that some bilinguals switch more often than others may affect the processing mechanisms they use to control lexical access in the two languages.
Appendix B

Language History and the Self-Assessed Proficiency for All Participants

The language history and self-assessed proficiency scores of the participants in all the experiments reported in the article are presented in Table B1. The mean age and standard deviation are given in years. The “L2 onset” refers to the mean age (in years) at which participants started learning Catalan, Basque, or English. “L2 use” refers to how long (in years) participants had or have been using L2 regularly. The same applies to the participants’ L3 and L4 scores. 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).

As is shown in Table B1, the Spanish–Basque early bilinguals (Experiment 1) and the Spanish–Catalan early bilinguals tested in Costa and Santesteban (2004b, Experiment 2) were quite similar in their language history and self-assessed L1 and L2 proficiency levels. However, between the Spanish–English late bilinguals (Experiment 2) and the Spanish–Catalan early bilinguals tested in Costa and Santesteban (2004b, Experiment 2), we can see a main difference of age of L2 acquisition (L2 onset), \( t(11) = 5.56, p < .001 \). Finally, a comparison of the two groups of Experiment 4 confirms the difference in L2 proficiency level between what we considered Spanish monolinguals and the group of Spanish–Catalan bilinguals (L2 onset): \( t(11) = 5.28, p < .001 \); L2 use, \( t(11) = 6.93, p < .001 \).

### Table B1

<table>
<thead>
<tr>
<th>Group</th>
<th>Language history</th>
<th>Self-assessed proficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L2 Onset Use</td>
<td>L3 Onset Use</td>
</tr>
<tr>
<td>Early Sp-Cat bilinguals (Exp. 2)*</td>
<td>Age</td>
<td>23 (3) 4 (2) 18 (3)</td>
</tr>
<tr>
<td>Early Sp-Bas bilinguals (Exp. 1)</td>
<td></td>
<td>22 (2) 3 (2) 19 (2)</td>
</tr>
<tr>
<td>Late Sp-Eng bilinguals (Exp. 1)</td>
<td></td>
<td>25 (6) 10 (4) 16 (4)</td>
</tr>
<tr>
<td>Sp-Cat bilinguals (Exp. 2)</td>
<td></td>
<td>21 (2) 5 (2) 16 (3)</td>
</tr>
<tr>
<td>Sp-Cat bilinguals (Exp. 3)</td>
<td></td>
<td>23 (5) 2 (3) 20 (5)</td>
</tr>
<tr>
<td>Sp monolinguals (Exp. 4)</td>
<td></td>
<td>24 (4) 10 (2)</td>
</tr>
<tr>
<td>Sp-Cat bilinguals (Exp. 4)</td>
<td></td>
<td>22 (2) 5 (2) 17 (3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>L1 L2 L3 L4 Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early Sp-Cat bilinguals (Exp. 2)*</td>
<td>3.9 (.2) 3.6 (.6) 0.3 (.4) (L1–L2)</td>
</tr>
<tr>
<td>Early Sp-Bas bilinguals (Exp. 1)</td>
<td>3.8 (.4) 3.6 (.5)</td>
</tr>
<tr>
<td>Late Sp-Eng bilinguals (Exp. 1)</td>
<td>4.0 (0) 3.7 (5) 0.3 (3) (L1–L2)</td>
</tr>
<tr>
<td>Sp-Cat bilinguals (Exp. 2)</td>
<td>4.0 (0) 3.6 (7) 2.1 (.7) 1.5 (.6) (L2–L3)</td>
</tr>
<tr>
<td>Sp-Cat bilinguals (Exp. 3)</td>
<td>3.9 (2) 3.8 (.1) 3.0 (.4) 1.6 (.7) 1.4 (.6) (L3–L4)</td>
</tr>
<tr>
<td>Sp monolinguals (Exp. 4)</td>
<td>3.8 (.1) 1.9 (.4) 1.9 (.7) 1.9 (.7) (L1–L2)</td>
</tr>
<tr>
<td>Sp-Cat bilinguals (Exp. 4)</td>
<td>4.0 (.4) 3.7 (.7) 0.3 (.4) (L1–L2)</td>
</tr>
</tbody>
</table>

Note. L = language; Sp = Spanish; Cat = Catalan; Bas = Basque; Exp. = experiment. Means with standard deviations in parentheses are reported.
* From Costa and Santesteban (2004b), Experiment 2.

Appendix C

Materials Used in Experiments 1, 2, and 4

In Experiments 1 (Group 2), 2, and 4, the 10 words marked with an asterisk were used (see Table C1). These materials were the same used in our previous study (Costa & Santesteban, 2004b). The materials marked with the symbol @ were used in Experiment 1 (Group 1). All the words used in these experiments were noncognates in the relevant languages. As the words carrot and hat are cognates in Spanish and Basque (zanahoria/azenarioa and sombrero/sonbreroa in Spanish and Basque, respectively), these pictures were substituted for the noncognate words owl and key. The Spanish and Basque names of the pictures were matched for number of phonemes: 4.7 and 5.2, respectively, \( t(9) = 1.26, p < .244 \). However, in Experiment 1 (Group 2), Spanish names were longer than English names (5.6 and 3.6 main number of phonemes, respectively), \( t(9) = 3.87, p < .004 \). In Experiment 2, the Catalan names of the pictures were longer than the English names as well (5.7 and 4.1 respectively), \( t(9) = 2.95, p < .016 \). In Experiment 4, care was taken to build “new-language” words with similar length (number of

(Appendixes continue)
phonemes: Spanish, 5.6; new language, 6.4), $t(9) = 0.92, p < .38$. To maximize the similarity of the learning process of a new language, we built some of the new-language words following phonotactic rules not allowed by the learners’ languages. For example, neither Spanish nor Catalan admit word-initial consonant clusters like “sl-” nor does their phonemic repertoire contain the allophonic geminate plosive forms of the phoneme “p” of the new-language word \textit{clappede} (pronounced like in the Italian word \textit{cappello [hat]}).

Table C1

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|}
\hline
Spanish name & Catalan name & English name & Basque name & New-language name \\
\hline
*Sombrero & Barret & Hat & & Inderut \\
*Zanahoria & Pastanaga & Carrot & & Pinca \\
*Manzana @ & Poma & Apple & Sagarra & Galv\v k \\
*Mesa @ & Taula & Table & Mahaia & Silto \\
*Perro @ & Gos & Dog & Txakurra & Slompuki \\
*Hoja @ & Fulla & Leaf & Hostoa & Pirte \\
*Cuchillo @ & Ganivet & Knife & Labana & Plavintu \\
*Ventana @ & Finestra & Window & Leihoa & Funo \\
*Lluvia @ & Pluja & Rain & Euria & Clappede \\
*Queso @ & Formatge & Cheese & Gazta & Guispemoke \\
Bu\’ho @ & & & & \\
Llave @ & & & & \\
\hline
\end{tabular}
\caption{Words Used in Experiments 1, 2, and 4}
\end{table}

Note. The 10 words marked with an asterisk were used in Experiments 1 (Group 2), 2, and 4. The materials marked with the symbol @ were used in Experiment 1 (Group 1).

Appendix D

Materials Used in Experiment 3

For Experiment 3, 8 out of 10 of the words used in the previous experiments were changed to avoid possible cognate interference from the participants’ dominant languages (Spanish and Catalan). The French and English names of the pictures were matched for number of phonemes: 4.6 and 4.0, respectively, $t(9) = 1.61, p < .140$. See Table D1.

Table D1

\begin{table}[h]
\centering
\begin{tabular}{|c|c|}
\hline
English name & French name \\
\hline
Hat & Chapeau \\
Dog & Chien \\
Bell & Cloche \\
Star & \textit{\^ {e}}toile \\
Glass & Verre \\
Leg & Jambe \\
House & Maison \\
Suitcase & Valise \\
Car & Voiture \\
Pencil & Crayon \\
\hline
\end{tabular}
\caption{Words Used in Experiment 3}
\end{table}

Received November 10, 2005
Revision received March 8, 2006
Accepted April 4, 2006

1074 COSTA, SANTESTEBAN, AND IVANOVA