

# Lexical Selection in Bilingual Speech Production Does Not Involve Language Suppression

Matthew Finkbeiner, Jorge Almeida, Niels Janssen, and Alfonso Caramazza  
Harvard University

The “hard problem” in bilingual lexical access arises when translation-equivalent lexical representations are activated to roughly equal levels and, thus, compete equally for lexical selection. The language suppression hypothesis (D. W. Green, 1998) solves this hard problem through the suppression of lexical representations in the nontarget language. Following from this proposal is the prediction that lexical selection should take longer on a language switch trial because the to-be-selected representation was just suppressed on the previous trial. Inconsistent with this prediction, participants took no longer to name pictures in their dominant language on language switch trials than they did on nonswitch trials. These findings indicate that nontarget lexical representations are not suppressed. The authors suggest that these results undermine the viability of the language suppression hypothesis as a possible solution to the hard problem in bilingual lexical access.

*Keywords:* bilingual, lexical selection, lexical access, language switching, speech production

When an individual is asked to name a picture or an object, the individual must first identify the object and retrieve the appropriate conceptual representation from memory. Following concept selection, the appropriate lexical representation must be retrieved. According to most models of speech production (e.g., Caramazza, 1997; Dell, 1986; Levelt, Roelofs, & Meyer, 1999), activation at the conceptual level cascades down onto the lexical level, causing several semantically related lexical representations to become activated simultaneously. Consequently, a lexical selection mechanism is needed to decide which lexical representation should be chosen for further processing. Lexical selection is necessary because lexical representations specify the phonological segments of the to-be-articulated word, and these segments must be assembled before articulation may commence. The process of translating conceptual information into articulated speech is known as *lexical access*, and although a great deal of progress has been made in the study of the dynamic and architectural properties of lexical access, there are still several aspects for which researchers have not yet reached a consensus. This is especially true in the study of bilingual speech production. Bilingual lexical access is potentially more complicated than monolingual lexical access because of the assumption that concept selection serves to activate two lexical representations to an equal extent in the bilingual mind. Theoretically speaking, this creates difficulty at the point of lexical selec-

tion because it is unclear how the lexical selection mechanism knows to select the target lexical node given that it and the nontarget translation-equivalent lexical node are activated to the same level. Finkbeiner, Gollan, and Caramazza (in press) have recently referred to this as the “hard problem” of bilingual lexical access. The nature of the hard problem is depicted in Figure 1.

In the case of monolingual picture naming, lexical selection is assumed to proceed rather straightforwardly because the target lexical item should always be the most highly activated item in the lexicon (Figure 1A). In the case of bilingual picture naming, though (Figure 1B), lexical selection is assumed to be more difficult because, as a result of their equivalent meaning, the target lexical node and its translation-equivalent lexical node are activated equally. This hard problem should be extensive in bilingual speakers because virtually each concept in the bilingual mind (especially concrete concepts; Tokowicz, Kroll, de Groot, & Van Hell, 2002) is associated with synonymous lexical nodes (see Peterson & Savoy, 1998, for a discussion of synonyms in monolingual lexical access). Furthermore, to the extent that models of bilingual lexical access predict that lexical selection should be difficult whenever the semantic system activates translation-equivalent lexical nodes to an equal degree, lexical selection should be the most difficult for highly proficient bilinguals.<sup>1</sup> Yet, highly proficient bilinguals rarely exhibit signs of lexical intrusions from their other language (Poullisse, 1997; Poullisse & Bon-

---

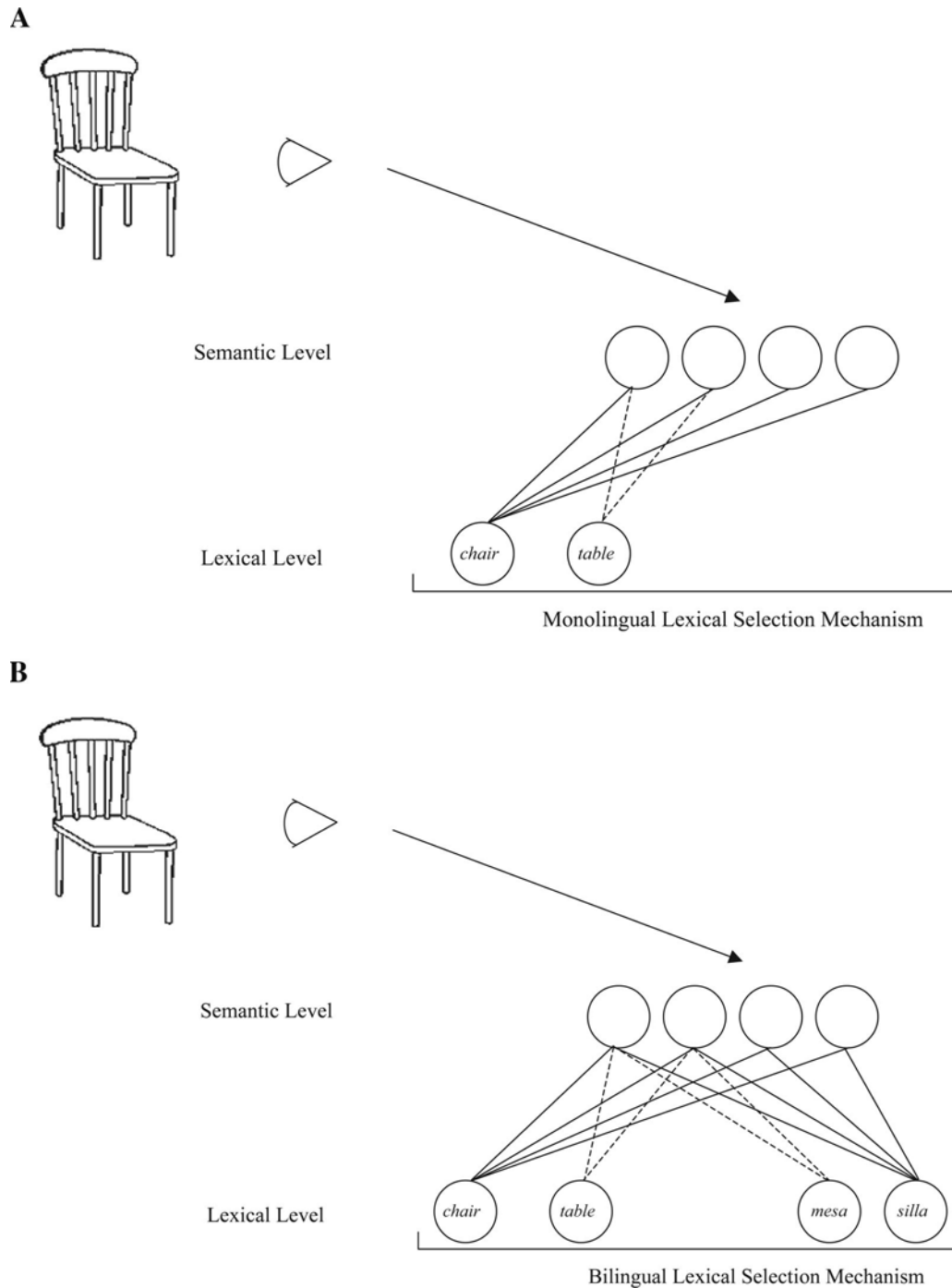
Matthew Finkbeiner, Jorge Almeida, Niels Janssen, and Alfonso Caramazza, Department of Psychology, Harvard University.

The research reported here was supported in part by National Institutes of Health Grant DC 04542-04 and a Fundação Calouste Gulbenkian’s research scholarship. We gratefully acknowledge David Green for helpful comments on an earlier version of this article. We also thank Adam Kundishora for help in testing participants and scoring responses.

Correspondence concerning this article should be addressed to Matthew Finkbeiner, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138. E-mail: msf@wjh.harvard.edu

---

<sup>1</sup> In less proficient bilinguals, it is reasonable to assume that the connections between the semantic system and the lexical system of the nondominant language (L2) are too weak for lexical representations to become activated highly enough to compete with their dominant language (L1) counterparts (cf. Kroll & Stewart, 1994; Potter, So, Von Eckardt, & Feldman, 1984). In highly proficient bilinguals, on the other hand, it is widely assumed that the semantic system activates lexical nodes in both languages to an equal degree.



*Figure 1.* Comparison of monolingual and bilingual lexical selection. In monolinguals, the target node is the most highly activated lexical node (in this case, *chair*), making selection relatively straightforward. In bilinguals, translation equivalents (in this case, *chair* and *silla*) become activated to equal degrees, requiring the lexical selection mechanism to decide which node is the target representation. We refer to this as the “hard problem” in bilingual lexical access.

gaerts, 1994). Clearly, then, either the assumption that the semantic system activates translation-equivalent lexical nodes to the same degree is wrong, or the assumption is correct and bilinguals have somehow overcome the difficulty that equally

activated translation equivalents present to the lexical selection mechanism. Both possibilities have been considered in the literature; in this article, we consider the possibility that the assumption giving rise to the hard problem is correct and

that bilinguals have developed an effective solution to the problem.

### Solving the Hard Problem in Bilingual Speech Production

Two distinct types of proposals have been made with respect to how bilinguals may go about solving the hard problem: The first proposes a lexical selection mechanism that considers the activation levels of lexical nodes only in the target language (cf. Costa, Miozzo, & Caramazza, 1999); the second proposes a mechanism that creates differences between the activation levels of lexical nodes in the target and nontarget languages. With respect to the latter approach, differential activation can be achieved either through activating lexical nodes in the target language more (i.e., the assumption giving rise to the hard problem is wrong; Finkbeiner et al., in press; La Heij, 2005; Poulisse & Bongaerts, 1994) or by reactively suppressing the lexical nodes in the nontarget language (Green, 1986, 1998). According to the latter approach, it is assumed that translation-equivalent lexical nodes become activated to roughly the same levels (i.e., the hard problem) and that the difficulty that this creates for the lexical selection mechanism is overcome through the subsequent (reactive) suppression of the nontarget language. The focus of this article is on the viability of language suppression as a possible solution to the hard problem.<sup>2</sup>

### *Evidence for Language Suppression*

The most direct source of empirical support for the language suppression hypothesis has come from language switching tasks and, in particular, from a study done by Meuter and Allport (1999). In Meuter and Allport's seminal study on language switching performance, participants were asked to name single digits (range = 1–9) presented on a computer monitor in either their dominant language (L1) or nondominant language (L2), depending on the color of the screen. In this experiment, there were four data points of primary interest. The first two were L1 and L2 naming latencies on nonswitch trials—trials in which the preceding setup trial was named in the same language as the target trial. The second two data points of interest were L1 and L2 naming latencies on switch trials—trials in which the setup and target trials were named in different languages. Meuter and Allport reported two effects of primary importance for our purposes here. The first was the main effect of switching. This refers to the increased naming latency for switch trials (averaged across L1 and L2 responses) relative to nonswitch trials. The second effect was the asymmetrical switch cost for L1 and L2 responses; critically, the switch cost for L1 responses was greater than the switch cost for L2 responses. This asymmetrical language switch cost has become the signature effect of language suppression because it follows directly from the language suppression hypothesis. This hypothesis is articulated best in Green's (1986, 1993, 1998) Inhibitory Control Model (ICM).

### *The ICM (Green, 1998)*

The ICM expands on the Norman and Shallice (1986) model by assuming that language processes and actions are under the same control mechanism. Specifically, according to this model of bilingual speech production, control is achieved through the implemen-

tation of *language task schemas*. On this model, each lexical representation is associated with a language tag (e.g., L1 or L2), and task schemas are said to exert control within the bilingual lexicon by activating and inhibiting lexical nodes on the basis of their language tags. Task schemas also exert control through the suppression of competing task schemas. For example, when the task goal is to name an object in L1, the L1 task schema assumes control of lexical selection processes by activating lexical representations with L1 tags and by suppressing the L2 task schema (which, in turn, serves to inhibit lexical representations with L2 tags). As such, the ICM specifies two loci of inhibition: inhibition of schemas that operate outside of the lexicon and inhibition of lexical representations within the bilingual lexicon. An important feature of the ICM is that inhibition is proposed to be reactive and proportional such that the more nontarget lexical representations become activated initially, the stronger those representations are then inhibited. Green (1998) made this point clearly when he stated that “because inhibition is reactive, more active lemmas will be more inhibited” (p. 74).

The findings reported by Meuter and Allport (1999) follow straightforwardly from the assumptions of the ICM. First, when switching from Language A to Language B, the inhibition of Language B must be overcome. Because it is assumed that overcoming inhibition incurs a cost, it naturally follows that some time will elapse before the Language B task schema can control lexical selection processes. Thus, picture-naming latencies should be longer on language switch trials than on nonswitch trials. Second, because L1 is the stronger language and, as Green (1998) argued, “because overcoming prior inhibition will be a function of the prior amount of suppression, it can be predicted that the cost of switching will be asymmetric. It will take longer to switch into a language which was more suppressed—for unbalanced bilinguals this will be L1, their dominant language” (p. 74). This last point highlights an important feature of this model. The critical factor that determines the magnitude of the switch cost is the strength of the nontarget language task schema on the previous trial. The stronger the nontarget language on Trial  $N - 1$ , the longer it will take to reactivate that language on Trial  $N$ . The findings reported by Meuter and Allport (1999) are perfectly consistent with these predictions.

Although the asymmetrical language switch cost is consistent with the predictions of the language suppression hypothesis, previous language switching experiments have conflated language membership with characteristics of the switching task that are known to give rise to asymmetrical switch costs in unilingual contexts. Thus, it has not yet been established that language switch costs, such as those reported by Meuter and Allport (1999), necessarily implicate language suppression. The purpose of the experiments reported here is to uncouple task-specific effects from those that can be attributed directly to language suppression in an effort to establish unambiguous effects of language suppression.

What characteristics of the switching task lend themselves to effects that can then be interpreted as effects of language suppression? There are two. One has to do with ease of response. For

<sup>2</sup> With this focus in mind, we defer discussion of the possibility that the assumptions giving rise to the hard problem are incorrect until the General Discussion.

example, Allport, Styles, and Hsieh (1994) had participants switch between naming the ink color of words and the words themselves, and they found that participants took longer to switch to the easier word-naming task than to the more difficult color-naming task. Although it is now clear that it does not always take longer to switch to the easier of two tasks (Monsell, Yeung, & Azuma, 2000), a greater switch cost for the easier task is generally found when the experimental procedures match those used by Meuter and Allport (1999; e.g., simultaneous cue–target onset). The second has to do with overlapping stimulus and response sets. Switch costs are reliably obtained with stimuli that afford two distinct responses (i.e., bivalent stimuli), but switch costs are difficult to obtain with nonoverlapping stimulus and response sets (i.e., univalent stimuli; but see Rogers & Monsell, 1995, Experiment 4). For example, in line with earlier reports in which no switch costs were observed for univalent stimuli (cf. Spector & Biedermann, 1976), Allport et al. (1994) found no evidence of switch costs in a task in which participants switched between color and numerosity naming of Stroop-like stimuli (e.g., said “red” to the word GREEN written in red ink and “three” to an array of three “4”s). Taken together, there are compelling reasons to think that certain stimulus characteristics can interact with the demands of the switching task to produce effects that can masquerade as effects of language suppression.

In an effort to disentangle switch costs that can be attributed to language suppression from those that can be attributed to particular characteristics of the stimuli, we held constant the language switch (L2–L1) and varied the stimulus type on the L1 switch trial. We accomplished this by comparing L1 switch costs for stimuli that were named only in L1 with those stimuli that were named in both L1 and L2. Following the standard terminology used in the task switching literature, we will refer to stimuli that were named only in L1 as “univalent” stimuli, and we will refer to stimuli that were named in L1 and L2 alike as “bivalent” stimuli.<sup>3</sup> Because univalent stimuli do not typically produce switch costs in unilingual experiments (see above), it is reasonable to attribute whatever switch costs univalent stimuli produce in a language switching context to the effects of language suppression. Remember, the language suppression hypothesis (Green, 1998) stipulates that language switch costs arise as a result of switching to a language that was inhibited on the previous trial. Furthermore, the stronger a language is, the more strongly it will be suppressed when it is the nontarget language. Hence, switching to the stronger language (L1) should always incur a cost, regardless of the stimulus’s valence on the current trial. In the following experiments, we tested this prediction of the language suppression hypothesis by comparing L1 switch costs for bivalent stimuli (i.e., stimuli that elicit both L1 and L2 responses) and univalent stimuli (i.e., stimuli that elicit L1 responses only).

In Experiment 1, we tested the strongest version of the language suppression hypothesis by introducing pictures (our univalent stimuli) into a digit-naming experiment similar to the one used by Meuter and Allport (1999). Digits were bivalent insofar as they were named in participants’ L1 and L2 depending on a color cue. Pictures were univalent stimuli insofar as they were named only in L1. Replicating Meuter and Allport, we observed an asymmetrical switch cost on the digit-naming trials, but inconsistent with the predictions of the language suppression hypothesis, we found no evidence of a language switch cost on the picture-naming (univa-

lent) trials. This finding indicates that the nontarget language is not suppressed as a whole when selecting lexical representations in the target language. In Experiment 2, we tested a weaker version of the language suppression hypothesis that stipulates that suppression operates over just those lexical nodes that compete with the target lexical node for lexical selection. We refer to this version of the hypothesis as the *lexical suppression hypothesis*. To test this version of the suppression hypothesis, we replaced the univalent stimuli used in Experiment 1 with dot patterns. It is important to note that the dot patterns elicited the same numerical responses as the digits but were univalent in the sense that they elicited L1 responses only. Once again, we obtained the asymmetrical language switch cost originally reported by Meuter and Allport on the digit-naming trials, but just as in Experiment 1, we found no evidence of language suppression on the trials in which responses were elicited by univalent dot patterns.

On the basis of the findings of Experiments 1 and 2, we conclude that there is no support for the proposal that lexical representations in the nontarget language are suppressed when selecting target-language lexical representations. Hence, our main conclusion in this article is that the language suppression hypothesis does not constitute a viable solution to the hard problem in bilingual lexical access. Taking our findings further, we consider in Experiment 3 the possibility that the restriction of the asymmetrical language switch cost to bivalent stimuli may have more to do with these stimuli affording one “easy” (or fast) and one “difficult” (or slow) response than one L1 and one L2 response. To test this possibility, we selected fast and slow word-naming stimuli (all in English) and found larger switch costs for fast words (e.g., “house”) than for slow words (e.g., “cottage”). Thus, a variable that is unavoidably conflated with language membership in language switching tasks (i.e., speed of response availability for L1 vs. L2 words) is sufficient to produce an asymmetrical switch cost within a task (i.e., word naming). Hence, our second conclusion in this article is that the asymmetrical language switch cost, though consistent with the predictions of the language suppression hypothesis, does not provide compelling support for this hypothesis.

## Experiment 1

To test the predictions of the language suppression hypothesis, we had L2 learners (unbalanced bilinguals) name digits in either their L1 or L2, depending on a color cue. In the same experiment, we had participants name pictures as well. The pictures were always named in the participants’ L1. Within the context of the experiment, the digits were considered bivalent stimuli insofar as they elicited L1 and L2 responses, and the pictures were considered univalent stimuli insofar as they elicited L1 responses only.

### Method

*Participants.* Sixteen undergraduate students at Harvard University participated for course credit or pay. All participants spoke English as their first language and rated their proficiency in English as a 7 on a 7-point

<sup>3</sup> It should be noted that the terms *univalent* and *bivalent* are accurate only within the context of the specific task demands. Technically, all stimuli were bivalent to our participants as they knew both L1 and L2 names for them.

scale, where 7 was perfect. All participants judged themselves to be relatively proficient in their second language ( $M = 5.1$  on a 7-point scale), and all had either lived or studied abroad in a country in which their second language was spoken as the primary language. Similar to Meuter and Allport (1999), we deliberately selected the languages spoken by our participants from a range of languages in order to limit systematic effects of cognate number names in particular language pairs. The languages were Chinese ( $n = 4$ ), French ( $n = 3$ ), German ( $n = 2$ ), Italian ( $n = 1$ ), Japanese ( $n = 1$ ), and Spanish ( $n = 5$ ).

**Materials.** In the digit-naming part of the experiment, the digits 1–9 were used. Single digits were presented in a 16-point Arial font and were superimposed on either a green or gray circle with a diameter subtending approximately  $10^\circ$  of visual angle on a computer monitor placed approximately 50 cm from the participant. For half of the participants, the green color cue was used to elicit an L1 naming response; for the other half, the gray cue was used to elicit an L1 naming response. Additionally, nine pictures were chosen from the International Picture Naming Project database (<http://crl.ucsd.edu/~aszekely/ipnp/>). The pictures had a name agreement of 100% in English. No other property was explicitly manipulated or controlled. The pictures were *bed, box, cake, dog, fish, flower, sun, tree, and turtle*.

**Design and procedure.** Participants were tested individually in a sound-attenuating and dimly lit booth. Participants were asked to name the digits and pictures as quickly as they could. The software program DMDX (Forster & Forster, 2003) was used to display the items and record participants' responses. Prior to the experiment proper, participants were told the meaning of the two color cues (e.g., green indicates L1, and gray indicates L2) and were asked to name each digit twice in the context of each color cue. In the practice session, trials were blocked according to response language, and the order of the blocks was counterbalanced across participants. Participants were also asked to name each picture twice prior to the experiment proper. They were told that pictures would randomly appear during the course of the experiment and that they were always to name the pictures in English (their L1). The pictures did not appear with a color cue.

Trials were arranged into groups of four. The first two trials in each quartet were filler or "setup" digit-naming trials and always appeared with the same color cue; the third trial was the target digit-naming trial. On nonswitch trials, the target digit-naming trial appeared with the same color cue as the setup trials; on switch trials, the target appeared with a different color cue from that used on the setup trials. Hence, each target digit-naming trial had a "run length" of two. In one half of the quartets, a picture was presented for naming on the fourth trial; in the remaining quartets, the fourth trial was omitted, and trials proceeded directly to the next quartet. The trials in each quartet were presented in a fixed order (three digit-naming trials followed by one picture-naming [or omitted] trial), but the quartets were presented in a different random order for each participant. This ensured that participants would not be able to reliably predict the trial or stimulus type.

In the digit-naming task, there were two experimental factors with two levels each: response language (L1 vs. L2) and trial type (nonswitch vs. switch). Each digit ( $N = 9$ ) appeared four times in each of the four experimental conditions for a total of 144 target trials. The picture-naming task also had two experimental factors with two levels each: preceding digit-naming trial type (nonswitch vs. switch) and response language on preceding digit-naming trial (L1 vs. L2). Each picture ( $N = 9$ ) appeared twice in each of the picture-naming experimental conditions, for a total of 72 target picture-naming trials.

Each trial began with a central fixation point that appeared in a black font on a white background for 500 ms. This was immediately followed by either a digit superimposed on a color cue or a picture. In the case of the digit-naming trials, the digit and the color cue appeared simultaneously. Pictures appeared without a color cue. Both digits and pictures stayed on the screen for 3 s or until the voice key was triggered by the participant's

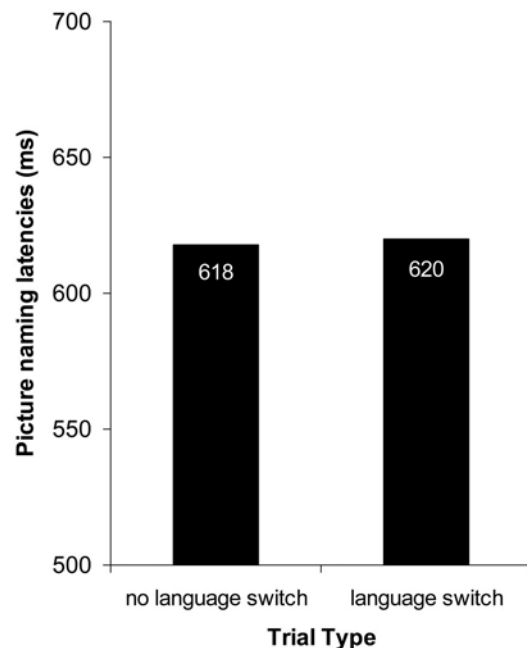
response, whichever came first. One full second elapsed between the triggering of the voice key and the onset of the next trial.

## Results

In this and all following experiments, errors, including incorrect responses and verbal disfluencies (e.g., stuttering, lip smacking, utterance repairs), as well as voice key failures, were excluded from the reaction time analysis. Outliers were treated by setting them equal to cutoffs established two standard deviation units above and below each participant's mean response latency, which was calculated across all items and naming conditions in each experiment. This procedure affected approximately 3.3% of the responses in the experiments reported below (3.3%, 3.6%, and 3.1% in Experiments 1, 2, and 3 respectively).

**Picture naming in L1 following L1 and L2 digit naming.** According to the language suppression hypothesis, it should take longer to name a picture in L1 on a language switch trial than it does on a nonswitch trial. The results from the picture-naming trials in this experiment do not confirm this prediction. Participants' picture-naming latencies were just as fast for pictures preceded by an L2 digit-naming trial ( $M = 620$ ) as they were for pictures preceded by an L1 digit-naming trial ( $M = 618$ ; all  $F_s < 1$ ). This result indicates that picture-naming latencies (in L1) were not modulated by the presence of a language switch between digit-naming and picture-naming trials (see Figure 2). The main effect of preceding digit-naming trial type (nonswitch vs. switch) did not reach significance, nor did the interaction between preceding trial type (nonswitch vs. switch) and preceding response type (L1 vs. L2; all  $p_s > .1$ ).

The error analysis on the picture-naming trials revealed very few errors but did reveal a small effect of language switching. The



**Figure 2.** Picture-naming latencies in Experiment 1 as a function of trial type (language nonswitch vs. language switch). Pictures were always named in participants' dominant language.

error rate in the nonswitch condition (preceding digit-naming trial was responded to in L1) was 0.02%, and the error rate in the language switch condition (preceding digit-naming trial was responded to in L2) was 1.6%. Although this difference was negligible numerically, it was statistically reliable,  $F_1(1, 15) = 7.5, p = .02$ ;  $F_2(1, 35) = 7.6, p < .01$ . Though this effect in the error analysis is consistent with the language suppression hypothesis, it does not replicate, as we will see in Experiment 2.

*Digit naming in L1 and L2.* Replicating earlier findings, a main effect of trial type was obtained, such that nonswitch trials were named faster than switch trials:  $F_1(1, 15) = 88.28, p < .01$ ;  $F_2(1, 35) = 390.37, p < .01$ . The main effect of response language was not significant in the subjects analysis but was in the items analysis:  $F_1(1, 15) = 2.03, p = .18$ ;  $F_2(1, 35) = 10.59, p < .01$ . Most important, the interaction between the two factors was significant (see Figure 3):  $F_1(1, 15) = 29.69, p < .01$ ;  $F_2(1, 35) = 49.62, p < .01$ . The nature of this interaction was just as Meuter and Allport (1999) reported: Participants took longer to switch to an L1 digit-naming trial than they did to switch to an L2 digit-naming trial.

The error analysis revealed a main effect of trial type only, with more errors occurring in the switch condition ( $M = 7.73$ ) than in the nonswitch condition ( $M = 1.92$ ):  $F_1(1, 15) = 16.01, p < .01$ ;  $F_2(1, 35) = 48.44, p < .01$ . None of the other effects were significant in the error analysis.

### Discussion

Two important results were obtained in Experiment 1. One was the replication of the asymmetrical language switch cost in the digit-naming task first reported by Meuter and Allport (1999). This, along with findings recently reported by Costa and Santesteban (2004; Experiment 1), confirms the robustness of the asymmetrical language switching cost with bivalent stimuli. The second result of importance was the finding that the asymmetrical lan-

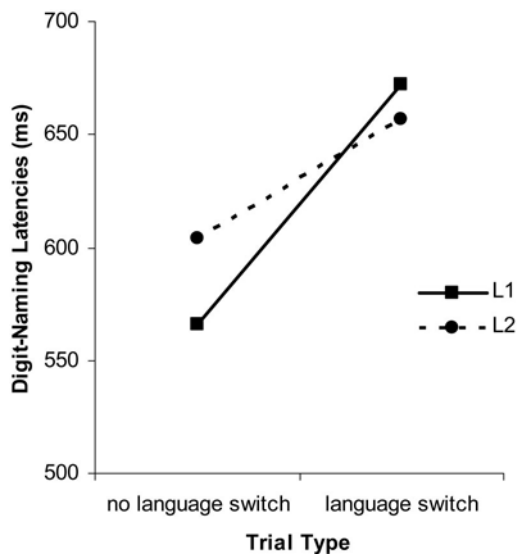


Figure 3. Digit-naming latencies in Experiment 1 as a function of trial type (nonswitch vs. switch) and language (L1 vs. L2). L1 = dominant language; L2 = nondominant language.

guage switching costs obtained with bivalent stimuli (digits) did not extend to univalent stimuli (pictures). Pictures, which were univalent insofar as they always elicited an L1 response, were named just as quickly following L1 digit-naming trials as they were following L2 digit-naming trials. This finding is at odds with the predictions of the language suppression hypothesis (Green, 1998). According to the language suppression hypothesis, lexical nodes in the nontarget language are suppressed by the language task schema controlling lexical selection processes in the target language. Hence, switching suddenly to the suppressed language should always incur a cost because of the time that is needed to overcome the inhibition on the previous trial. Furthermore, the language suppression hypothesis predicts that it should take longer for unbalanced bilinguals to switch into the dominant language because the inhibition of the language task schema controlling L1 production is greater and more difficult to overcome (Green, 1998). Yet, Experiment 1 revealed that the language suppression hypothesis makes the wrong predictions for univalent stimuli because these stimuli were named just as fast on trials involving a language switch as they were on trials not involving a language switch. Following from this, it would appear that the ability to select L1 lexical nodes for production was completely unaffected by the response language on the previous trial. It is unclear how the language suppression hypothesis could accommodate these findings. In Experiment 2, we tested a modified version of the language suppression hypothesis that proposes that suppression operates over just those nontarget-language lexical nodes that compete with target nodes for selection.

### Experiment 2

Although the results of Experiment 1 challenge the ICM (Green, 1998), especially the claim that suppression operates over competing language task schemas, they do not necessarily challenge the possibility that reactive suppression operates over competing lexical nodes only. As such, we may formulate a modified version of the suppression hypothesis; we will refer to this modified hypothesis as the lexical suppression hypothesis. According to the lexical suppression hypothesis, suppression operates over just those lexical nodes in the nontarget language that become as highly activated as target lexical nodes—that is, suppression targets just those nodes that create difficulties for the bilingual lexical selection mechanism.<sup>4</sup> According to this version of the suppression hypothesis, switching languages would not necessarily incur a cost unless the target on the switch trial is semantically related to the target on the previous trial. For example, when an individual is asked to name the stimulus “1” in Spanish (“uno”), it is assumed that the lexical nodes *one* and *uno* will become equally activated, thus requiring the nontarget lexical node *one* to be suppressed so that the lexical node *uno* may be selected. Insofar as the nontarget lexical nodes *two* and *three* and *nine* also become activated and vie for selection, they too will be suppressed. On the assumption that suppression is reactive and proportional, the nontarget lexical node

<sup>4</sup> In a recent article, de Groot and Christoffels (in press) distinguished between “global” and “local” levels of control, a distinction that shares similarities with the difference between the language and lexical suppression hypotheses.

*one* will be suppressed more strongly than *two*, which, assuming an effect of distance, will be suppressed more strongly than *nine*. Also, in unbalanced bilinguals, the L1 lexical node, in this case *one*, will be suppressed more strongly when the appropriate response is “uno” than the L2 lexical node *uno* will be suppressed when the appropriate response is “one.” It is important to note that this version of the suppression hypothesis would not predict a switch cost when switching from “uno” to “turtle” because the stimulus on the previous trial (“1”) would not have activated the lexical node *turtle*, and thus, *turtle* would not have been suppressed. In sum, the lexical suppression hypothesis, by restricting suppression to just those nontarget lexical nodes that compete for lexical selection, constitutes a possible solution to the hard problem in bilingual lexical access, is able to predict an asymmetrical switch cost in tasks involving switching between L1 and L2 responses for semantically related items, and does not predict switch costs when switching between semantically unrelated stimuli (which was the case in Experiment 1). To test the lexical suppression hypothesis, we took advantage of the fact that lexical nodes (e.g., *three*) become activated (and, thus, subject to suppression) by a variety of different physical stimuli (e.g., “3,” “three,” “1 + 2,” “3 × 1,” and “.:”). Because, according to the lexical suppression hypothesis, suppression operates over lexical representations that are invariantly activated by a range of different physical stimuli, switching languages (e.g., L2 to L1) and stimulus types (e.g., digits to dot patterns) should produce a switch cost that is comparable (if not greater) in magnitude to the cost of switching languages alone. To test this prediction of the suppression hypothesis, we had bilinguals name bivalent stimuli (digits) in both languages and univalent stimuli (dot patterns) in their L1 only. Crucially, both the digits and the dot patterns activated the same set of lexical representations (i.e., lexical nodes one through nine).

### Method

**Participants.** Sixteen undergraduate students at Harvard University participated for course credit or pay. All participants spoke English as their first language and rated their proficiency in English as a 7 on a 7-point scale, where 7 was perfect. Again, all participants judged themselves to be relatively proficient in their second language ( $M = 4.8$  on a 7-point scale), and 13 of the 16 had either lived or studied abroad in a country in which their second language was spoken as the primary language. Once again, to limit possible systematic effects of cognate number names in particular language pairs, we deliberately selected the languages spoken by our participants from a range of languages. The languages were Chinese ( $n = 4$ ), French ( $n = 2$ ), German ( $n = 3$ ), Japanese ( $n = 2$ ), Irish Gaelic ( $n = 1$ ), and Spanish ( $n = 4$ ).

**Materials.** The materials for the digit-naming task were identical to Experiment 1. The picture-naming stimuli, on the other hand, were different. In Experiment 2, the pictures were dot patterns: domino-like shapes consisting of a horizontally oriented rectangle with a vertical dividing line that separated a number of dots (range = 1–3) on the left from a number of dots on the right (range = 1–3). All combinations of one to three dots on each side were depicted, for a total of nine dot patterns. The dot patterns did not appear with a color cue.

**Design and procedure.** The design and procedure were identical to those of Experiment 1 except for the following changes to the pattern-naming task. Because of the concern that the stimulus–response mappings for the dot patterns would be too similar to those of the digits, we had participants verbally indicate the product of the dots on the left and right. To ensure that participants produced their responses automatically, we gave them an extended practice session before the experiment proper. In

the practice session, participants named each dot pattern ( $N = 9$ ) six times for a total of 54 practice pattern-naming trials. Additionally, 36 nonswitch target pattern-naming trials were added to the experiment proper. In Experiment 1, all picture-naming trials were “task switch” trials—that is, all picture-naming trials were preceded by a digit-naming trial. The inclusion of nonswitch pattern-naming trials in the present experiment allowed us to determine the cost of switching from digit to pattern naming and to get a more sensitive measure (compared with Experiment 1) of whether the task-switching cost was modulated by the presence of a language switch.

### Results

**Picture naming in L1 following L1 and L2 digit naming.** The inclusion of nonswitch pattern-naming trials (i.e., target trials were preceded by other pattern-naming trials) allowed us to determine the cost of switching from digit naming to pattern naming. A repeated measures analysis of variance of the factor trial type (nonswitch, switch + language switch, switch – language switch) revealed a main effect:  $F_1(2, 30) = 10.37, p < .01$ ;  $F_2(2, 70) = 3.97, p = .02$ . Post hoc *t* tests revealed that the cost of switching from digit naming to picture naming was reliable both when a language switch was and was not involved (all  $ps < .05$ ). Most important for our purposes here, the digit- to picture-naming switch cost was not modulated by the presence of a language switch (all  $ps > .09$ ). It is interesting to note that the numerical (nonsignificant) difference between the language switch trials and the no-language switch trials revealed that the language switch trials were faster (see Figure 4). Although this numerical difference was not statistically reliable, it is quite remarkable nevertheless because it shows very clearly that the L1 lexical nodes corresponding to numerals 1–9 could not have been suppressed when naming digits in L2—otherwise there would be at least a trend toward a greater cost on the language switch trials.

The mean error rates were 5.2%, 5.2%, and 3.1% in the non-switch, switch (no language switch), and switch (language switch) conditions, respectively. The error rates on the picture-naming trials were noticeably larger in this experiment ( $M = 4.5\%$ ) compared with those in Experiment 1 ( $M = 0.8\%$ ), but repeated measures analysis revealed no reliable effect of trial type (all  $F_s < 1$ ). The lack of an error effect in this experiment suggests that the statistical effect in the error analysis in Experiment 1 was spurious and uninformative, especially as the magnitude of the effect in Experiment 1 was so small numerically (0.8%).

**Digit naming in L1 and L2.** Just as in Experiment 1, a main effect of trial type (nonswitch vs. switch) was obtained, with nonswitch digit-naming trials being named much faster than switch trials:  $F_1(1, 15) = 70.14, p < .01$ ;  $F_2(1, 35) = 244.88, p < .01$ . Unlike Experiment 1, an effect of response language was also obtained, with digits being named in L1 faster than in L2:  $F_1(1, 15) = 17.65, p < .01$ ;  $F_2(1, 35) = 109.32, p < .01$ . Most important, the interaction between the two factors was once again significant:  $F_1(1, 15) = 29.05, p < .01$ ;  $F_2(1, 35) = 49.46, p < .01$ . As is shown in Figure 5, the nature of this interaction was such that participants took longer to switch to their L1 ( $M = 98$  ms) than to their L2 ( $M = 44$  ms).

The error analysis revealed a main effect of trial type, with more errors occurring in the switch condition ( $M = 6.9$ ) than in the nonswitch condition ( $M = 2.5$ ):  $F_1(1, 15) = 10.26, p < .01$ ;  $F_2(1, 35) = 28.79, p < .01$ . None of the other effects were significant in the error analysis.

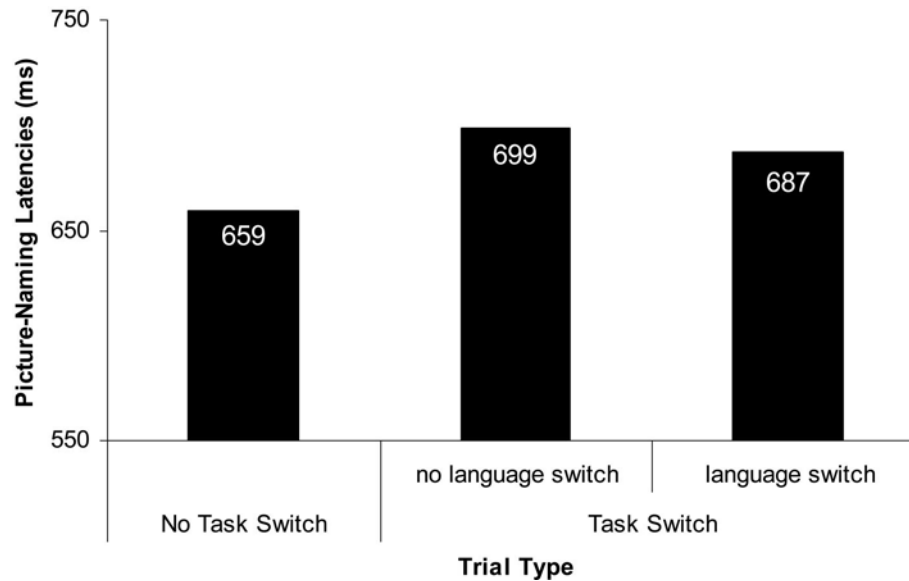


Figure 4. Picture-naming latencies in Experiment 2. All pictures were named in participants' dominant language. Task switch trials were preceded by a digit-naming trial; nontask switch trials were preceded by a picture-naming trial. Language-switch trials were preceded by a digit-naming trial in participants' nondominant language.

### Discussion

Three results of interest were obtained in Experiment 2. First, we replicated once again the asymmetrical switch cost originally observed by Meuter and Allport (1999) in the digit-naming task. This asymmetry, indexed by a larger switch cost for L1 responses,

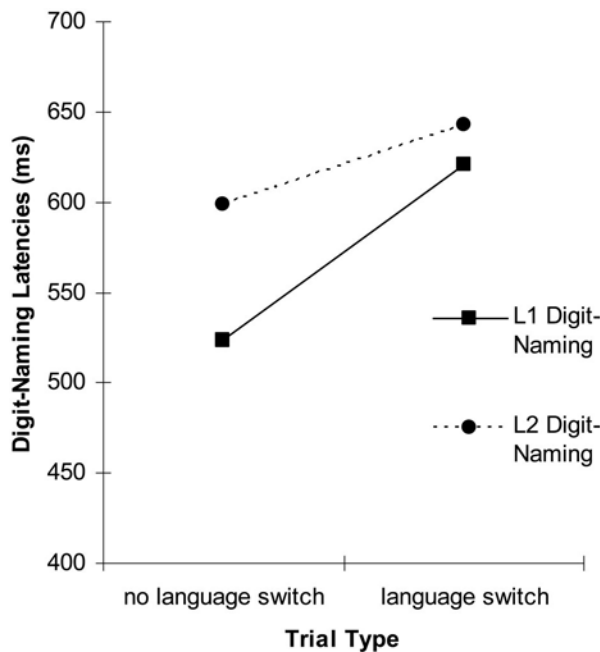


Figure 5. Digit-naming latencies in Experiment 2 as a function of trial type (nonswitch vs. switch) and language (L1 vs. L2). L1 = dominant language; L2 = nondominant language.

is taken to be a signature effect of language suppression. Second, just as in Experiment 1, we found no evidence of this signature effect of language suppression when participants switched from digit-naming trials to pattern-naming trials. This was the case despite the fact that digits and dot patterns elicited L1 responses from the same numerical set. As such, just as the results of Experiment 1 refute the strongest version of the language suppression hypothesis, the results of Experiment 2 refute the predictions of the weaker lexical suppression hypothesis. According to the lexical suppression hypothesis, suppression operates over just those nontarget lexical nodes that vie for selection. Thus, the reason we did not observe effects of language suppression in the picture-naming task in Experiment 1 was because the lexical nodes corresponding to the pictures were not activated on the preceding digit-naming trials and, hence, were not suppressed. In Experiment 2, though, the L1 lexical nodes corresponding to the dot patterns and the digits were the same. Thus, it logically follows that if the cost of switching from, say, L2 "uno" to L1 "six" in the digit-naming task is due to the lexical node *six* being suppressed when selecting the lexical node *uno*, then this same cost should be observed when switching from "uno" elicited by a digit to "six" elicited by a pattern of dots. Yet, the so-called language switch costs were restricted to Arabic numerals only. These findings indicate that the signature effect of language suppression is not due to suppression of lexical representations in the nontarget language, but, rather, that it is an artifact of using bivalent stimuli that elicit both L1 and L2 responses.

The third result obtained in Experiment 2 was a significant cost of switching from digit to picture naming. Because it is not the focus of this study to provide an account of task switching, we will not venture an explanation of this effect except to point out that it could be attributed to *task set inertia* (Allport et al., 1994; Meuter & Allport, 1999), or the time that is needed for the individual to

reconfigure task goals (Monsell, 2003; Rogers & Monsell, 1995). Regardless, it is clear that the digit- to pattern-naming switch was not modulated by the presence of a switch in the response language. This undermines a central tenet of the ICM, which seeks to solve the presumed difficulties of bilingual lexical selection by stipulating that lexical representations in the nontarget language are suppressed when selecting target-language lexical nodes.

The findings of Experiments 1 and 2 reveal that the language suppression hypothesis makes the wrong predictions in the language switching task—the very task that has yielded the signature effect of language suppression. Nevertheless, the asymmetrical language switch cost in digit naming is highly robust and deserves further mention. Although it is not the aim of this article to provide a theoretical account of how individuals perform switching tasks, the findings of Experiment 1 and 2 do suggest that this supposed signature effect of language suppression may simply be an artifact of switching between easy and difficult responses elicited by bivalent stimuli (cf. Allport et al., 1994; Allport & Wylie, 1999; Yeung & Monsell, 2003a, 2003b). This is an interesting possibility insofar as it suggests that the asymmetrical language switch cost may have nothing to do with language switching per se. We briefly consider this possibility in the following experiment in which all of the responses are in the participants' L1.

A distinguishing characteristic between L1 and L2 is the speed or ease with which L1 and L2 responses are produced. The rationale for Experiment 3 was to see whether manipulating “ease of processing” variables in a within-language experiment could produce the same asymmetrical switch cost observed in language switching experiments. It is important to note that ease of processing is a continuous variable. This is an important feature because whereas it is reasonable to think that L1 and L2 words may be tagged differently on the basis of their language membership and, thus, subject to suppression independently of the other, it is unlikely that words on the fast end of the response continuum are tagged differently than words on the slow end. Hence, an asymmetrical switch cost between fast and slow words, where fast words exhibit a larger switch cost than slow words, would constitute converging evidence that something besides language membership and language suppression is responsible for the asymmetrical switch cost in digit naming.

### Experiment 3

In this experiment, we replaced the dichotomous and plausibly taggable feature “language membership” with a continuous and untaggable feature “speed of response availability” to see whether differences in ease (speed) of processing, which is unavoidably confounded in language switching experiments, may not be sufficient to produce the asymmetrical switch cost reported by Meuter and Allport (1999) and in Experiments 1 and 2 above. We did this by having participants switch back and forth between naming the ink color of “fast” and “slow” words and the words themselves (cf. Allport et al., 1994).

#### Method

*Participants.* Thirty-two undergraduates (16 in Experiment 3a and 16 in Experiment 3b) at Harvard University participated either for course credit or pay. All participants were native speakers of English and had normal or corrected-to-normal vision.

*Materials.* In an effort to create an experimental context similar to that of a language switching experiment, we selected ten concepts with two possible responses (e.g., *dog–puppy*, *house–cottage*, *stone–pebble*) where, it is important to note, one of the responses could be designated as the “fast response” and the other could be designated the “slow response.” This designation was based on differences in lexical properties known to affect response availability (e.g., frequency, length, and number of semantic senses). The fast items had a much higher frequency ( $M = 114.5$ , CELEX; Baayen, Piepenbrock, & Gulikers, 1995) than the slow items ( $M = 34.9$ ), and the fast items had fewer letters ( $M = 4.5$ ) than the slow items ( $M = 6$ ). Additionally, the fast items had many more senses ( $M = 12.7$ ) than their counterparts ( $M = 1.3$ ).<sup>5</sup> In addition to the targets, 20 unrelated words were selected to be used on setup trials. None of the words were color words, and none of the words had an obvious canonical color (e.g., *blood*; see Appendix).

*Design and procedure.* The design and procedure were very similar to Experiments 1 and 2, but with the following differences. All words appeared in one of four different colors (red, green, yellow, or blue), and participants were asked to alternate between naming the color of the word and the word itself, depending on a background color cue. When the words appeared on a black background, participants were told to name the color of the word; when words appeared on a light gray background, participants were instructed to name the word itself. Similar to Experiments 1 and 2, items were arranged into groups of three. The first two items in each triad were filler or setup items, and the third item was the target trial. On nonswitch trials, all three items in a triad were of the same response type (color or word naming). On switch trials, the first and second items in the triad were of one response type, either color or word naming, and the target item was then from the other response type. For example, when a word-naming target appeared in the switch condition, participants named the color of the first and second items in the triad and then switched to word naming on the third (target) trial. Because we were interested in reproducing the experimental conditions of the language switching paradigm, where switches are from L1 (fast) to L2 (slow) or vice versa, we arranged the triads so that fast targets followed slow setup trials and vice versa. Also, to test further the possibility that suppression operates only over those lexical nodes that vie for selection, as opposed to tasks in general (see introduction to Experiment 2), we compared switch costs in Experiment 3a, where the setup and target trials were related in meaning (e.g., “stone” preceded “pebble”), with those of Experiment 3b, where the setup and target trials were not related (e.g., “stone” preceded “kitten”). Each target ( $N = 20$ ) appeared in each of the four different naming conditions (switch and nonswitch color naming and switch and nonswitch word naming), requiring 80 different triads to be constructed for a total of 240 trials. Item construction and presentation were identical to the previous experiments.

#### Results

An omnibus repeated measures analysis of variance included the factors Experiment (3a vs. 3b), Response Type (color naming vs. word naming), Trial Type (nonswitch vs. switch), and Word Type (fast vs. slow). Experiment was a between-subjects factor; the remaining factors were within-subjects factors. Not surprisingly, this analysis revealed that switch trials were reliably slower than nonswitch trials,  $F_1(1, 30) = 134.18, p < .01$ ;  $F_2(1, 18) = 257.27, p < .01$ ; and color-naming trials were significantly slower than

<sup>5</sup> In a recent article, Finkbeiner, Forster, Nicol, and Nakamura (2004) have suggested that L1 words have more senses than L2 words. See also Finkbeiner (2003) and Rodd, Gaskell, and Marslen-Wilson (2002), who have found that many-sense words are responded to faster than few-sense words. Sense counts were taken from Wordnet (Fellbaum, 1998), which is available on the Internet (<http://wordnet.princeton.edu/>).

word-naming trials,  $F_1(1, 30) = 71.73, p < .01$ ;  $F_2(1, 18) = 44.91, p < .01$ . There was no main effect of experiment or word type (all  $F_s < 1$ ). The three-way interaction between trial type, response type, and word type was significant— $F_1(1, 30) = 5.98, p = .02$ ;  $F_2(1, 18) = 7.24, p = .02$ —as were the two-way interactions between trial type and response type— $F_1(1, 30) = 13.94, p < .01$ ;  $F_2(1, 18) = 9.77, p < .01$ —and between trial type and word type— $F_1(1, 30) = 18.19, p < .01$ ;  $F_2(1, 18) = 9.57, p < .01$ . The nature of these interactions is clear in Figure 6. Essentially, the three-way interaction reveals that the interaction between word type (fast vs. slow) and trial type (nonswitch vs. switch) is much greater in word naming than in color naming. The interaction between trial type and response type (color naming vs. word naming) indicates that it takes longer to switch to word naming than to color naming. The interaction between trial type and word type indicates that the switch cost for fast words was greater than it was for slow words.

The error analysis with the same factors revealed a main effect of switching— $F_1(1, 30) = 8.85, p < .01$ ;  $F_2(1, 18) = 11.19, p < .01$ —with more errors occurring on switch trials ( $M = 7.58$ ) than on nonswitch trials ( $M = 4.86$ ), but no other main effects or interactions.

Because we were interested in the effects of the speed of response availability variable on switching, and because this variable has the strongest effect in word naming (with only marginal effects in color naming), our analysis of interest involved the word-naming trials. In an analysis of word-naming performance, the factors Experiment (3a vs. 3b), Trial Type (nonswitch vs. switch), and Word Type (fast vs. slow words) were included. Again, Experiment was a between-subjects factor, whereas the remaining factors were within-subjects factors. Once again, a main effect of trial type was obtained, such that nonswitch word-naming trials were named much faster than switch trials— $F_1(1, 30) = 97.24, p < .01$ ;  $F_2(1, 9) = 143.98, p < .01$ —but no main effect of word type was obtained (all  $F_s < 1$ ). The three-way interaction

was not significant (all  $F_s < 1$ ), suggesting that the switch costs were not modulated by semantic relatedness between setup and target. Crucially, the two-way interaction between trial type and word type was very robust:  $F_1(1, 30) = 26.28, p < .01$ ;  $F_2(1, 9) = 21.20, p < .01$ . As is clear in Figure 6, fast-response items exhibited a greater switch cost than slow-response items. Post hoc analyses revealed that fast words were named faster than slow words on nonswitch trials,  $F_1(1, 30) = 5.81, p = .02$ , and slower than slow words on switch trials,  $F_1(1, 30) = 12.98, p < .01$ . Follow-up analyses confirmed that the two-way interaction between trial type and word type was also significant when each experiment was analyzed separately: Experiment 3a,  $F_1(1, 15) = 10.63, p < .01$ ; Experiment 3b,  $F_1(1, 15) = 16.85, p < .01$ .

The error analysis revealed a main effect of trial type only. The mean error rate in the nonswitch condition ( $M = 4.7\%$ ) was significantly less than in the switch condition ( $M = 8.4\%$ ):  $F_1(1, 30) = 4.9, p = .03$ ;  $F_2(1, 9) = 16.11, p < .01$ . No other main effects or interactions reached significance in the error analysis.

### Discussion

There are several results of interest in Experiment 3. First, the cost of switching was greater for fast words than it was for slow words. Second, this asymmetrical switch cost for fast versus slow words was much greater in word naming than it was in color naming. Presumably, this is because the speed with which a word-naming response becomes available for production affects word-naming latencies much more than color-naming latencies. Third, it took participants longer to switch to a word-naming trial (139 ms) than it did to switch to a color-naming trial (78 ms). This latter finding replicates the findings reported by Allport et al. (1994) and extends those findings by showing that an asymmetrical switch cost can be obtained in a color-word switching paradigm even when the targets are not color words.

The most important result in Experiment 3 is the finding that in the word-naming task there was a greater switch cost for fast words (177 ms) than there was for slow words (101 ms). This result confirms that a continuous variable—namely, speed of response availability—is sufficient to produce an asymmetrical switch cost in tasks that are sensitive to this variable. It is interesting to note that if speed of response availability is a critical variable in obtaining the asymmetrical switch cost, then one may imagine that the cost of switching should be correlated with response speed in the nonswitch condition—and it is. In an analysis in which the switch cost for each item was correlated with its corresponding item mean in the nonswitch condition, a significant correlation was obtained:  $r(18) = -0.53, p = .01$ . Essentially, the faster an item is named on nonswitch trials, the longer it takes to switch to that item on switch trials. This finding has important implications with respect to how the asymmetrical language switching cost should be interpreted because, apparently, ease (or speed) of response availability is sufficient to produce an asymmetrical switch cost. Because the speed of response availability difference is also present in the language switching tasks between L1 and L2 responses, there is reason to think that the combination of factors that leads to differences in response availability may be the critical source of the asymmetrical language switch cost—not language suppression.

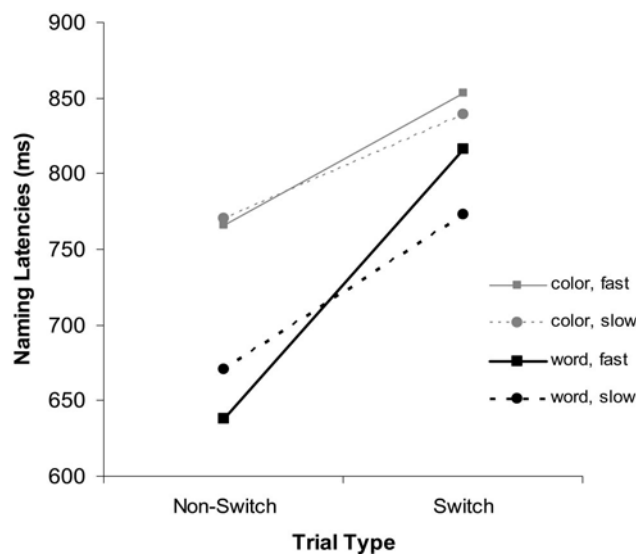


Figure 6. Color- and word-naming latencies as a function of word type (fast vs. slow) and trial type (switch vs. nonswitch) in Experiment 3 (collapsed across Experiments 3a and 3b).

Of course, if the ICM (Green, 1998) can account for the findings of Experiment 3, then it cannot be undermined by them. We consider two possible ways in which the ICM may be able to account for the findings of Experiment 3. For example, one could argue that the strength of the competing word-naming task on color-naming trials varies as a function of the processing difficulty of the word stimulus. That is, the more prepotent (readily available) a word-naming response is on a color-naming trial, the more the nontarget word-naming task will need to be suppressed, and hence, the longer it will take to produce a word-naming response on the subsequent trial. But this possible explanation is undermined by the results. First, looking at Figure 6, it is clear that color-naming latencies (in the nonswitch condition) were unaffected by the prepotency of the word-naming response. Participants named the color of fast and slow letter strings (e.g., "house" and "cottage") equally fast (means were 766 and 770, respectively). Furthermore, to the extent that the word-naming task might have been suppressed more when naming the ink color of "house" versus "cottage," one would expect longer word-naming latencies following words like "house." Yet, the results indicate the opposite. The slowest word-naming latencies on switch trials were for fast words, and fast word-naming targets always followed color naming of slow words (e.g., "cottage").

An additional aspect of the results of Experiment 3 that is problematic for the suppression hypothesis is the crossover interaction obtained for word-naming trials. On nonswitch trials, fast words were named faster than slow words, but on switch trials fast words were named slower than slow words. As mentioned above, post hoc analyses revealed that the differences between fast and slow words were significant on nonswitch and switch trials. A suppression account has difficulty explaining this finding. Take, for example, a fast word that has a resting level of five units of activation and a slow word that has only two units of activation. Insofar as the rationale for positing a suppression mechanism to control lexical selection in the color-word switching task is to ensure that the fast words are not any more competitive than the slow words when selecting a color name, fast words should be suppressed down to the level of the slow words (i.e., two units of activation). Yet, in order for a suppression hypothesis to account for the crossover interaction, it would have to stipulate that fast words were suppressed independently of, and to a greater extent than, the slow words. Presently, it is not clear how this stipulation could be motivated. The purpose of the suppression mechanism is to prevent competition from the word-naming responses when naming colors, and it does not follow from this original motivation to suppress nodes with five units of activation to zero, while allowing those with two units of activation to maintain their two units of activation.

We conclude that the language suppression hypothesis is undermined by the findings reported in Experiment 3. The findings of this experiment reveal that the continuous variable speed of response availability, which is also present in the L1–L2 distinction, is sufficient to produce an asymmetrical switch cost in a unilingual switching task. These findings suggest that it is not necessary to posit a suppression mechanism that operates over lexical representations on the basis of their language membership to account for an asymmetrical language switch cost. Furthermore, the findings reported in Experiments 1 and 2 indicate that a language switch cost is not obtained in the very conditions for which the

language suppression hypothesis clearly predicts a switch cost. Below we discuss the implications of these findings for models of bilingual lexical access.

## General Discussion

We have reported three experiments designed to test a central assumption of the language suppression hypothesis (Green, 1998), which is that bilinguals ensure successful selection of target-language lexical representations by suppressing nontarget-language lexical representations. The asymmetrical language switch cost, characterized by the finding that it takes longer to switch into the dominant language than it does to switch into the less dominant language, has been taken as the signature effect of language suppression. The findings reported in this article call into question the accepted interpretation of the asymmetrical language switch cost.

Three main findings have been reported in this article. First, we have successfully replicated the original asymmetrical language switch cost reported by Meuter and Allport (1999) in a digit-naming task in which unbalanced bilingual participants named digits (1–9) in either their L1 or L2 depending on the color of the computer screen. Second, in the same experiments, we found that the asymmetrical language switch cost is limited to bivalent stimuli. We found no evidence to suggest that language suppression extends onto trials requiring an L1 response when the stimuli eliciting those responses were consistently named in L1. This was true even when the lexical selection mechanism had to select from amongst the same set of representations when naming univalent and bivalent stimuli (Experiment 2). Third, we found that we were able to replicate the asymmetrical switch cost in a unilingual experiment in which stimuli were divided into fast- and slow-response items. The purpose of the unilingual experiment was to test the possibility that the continuous variable speed of response availability, which is shared by the L1–L2 distinction but is continuous and thus not readily available for suppression, may be sufficient to produce the asymmetrical switch cost. This was found to be the case.

### *Language Suppression Hypothesis*

The language suppression hypothesis was developed as a solution to the hard problem in bilingual speech production. This hard problem stems from two widely held assumptions in the bilingual speech production literature. The first of these is that the intention to say "dog," for example, activates two translation-equivalent lexical representations to an equal degree in the (proficient) bilingual mind (e.g., *dog* and *perro*). The second assumption is that lexical selection is competitive and that lexical selection becomes more difficult as the difference in activation levels between lexical nodes decreases (Levelt et al., 1999; Roelofs, 1992). In the case of our example, then, it should be very difficult for Spanish–English bilinguals to say "dog" because the lexical nodes *dog* and *perro*, by virtue of being activated to equal degrees, will compete fiercely with each other for selection. The language suppression hypothesis proposes a solution to this problem by tagging lexical nodes according to language membership and suppressing those nodes that do not correspond to the target language. According to this hypothesis, the intention of our hypothetical bilingual to say "dog"

would not only serve to activate *dog* but would result in the suppression of its translation equivalent *perro* as well. In this way, the lexical selection mechanism is able to proceed in much the same way as it would in the monolingual mind.

According to the ICM (Green, 1998), inhibitory processes operate both within and outside of the bilingual lexicon. Green (1998) proposed that each language system is associated with a language schema and that these schemas, when activated, suppress both competing language schemas (outside of the lexicon) and nontarget-language lexical nodes within the lexicon. For a language schema to control selection processes in Language A, the schema controlling selection processes in Language B must be suppressed. Furthermore, the stronger a particular language is, the more strongly its language schema must be suppressed when it is the nontarget language. The asymmetrical language switch cost follows directly from this formulation of inhibitory processes. In the language switching task, individuals take longer to name a stimulus on a language switch trial because suppression on the previous trial of the now-relevant language schema is said to persist onto the current trial. The asymmetrical switch cost, characterized by a greater switch cost for L1 than for L2, is attributed to the L1 language schema being suppressed more severely than the L2 schema (by virtue of the L1 schema being more dominant in unbalanced bilinguals) and the extra time that is needed to overcome this greater amount of suppression. Accordingly, the asymmetrical language switch cost is said to constitute a “signature effect” of language suppression.

Does the language suppression hypothesis correctly predict performance? The short answer is no. In Experiments 1 and 2, participants exhibited an effect that masqueraded as an effect of language suppression (the asymmetrical language switch cost). Though this effect is consistent with the predictions of the language suppression hypothesis, so would be a switch cost for stimuli named in just one language. In Experiments 1 and 2, no language switch cost was obtained for pictures that were named only in L1. This was true even when the responses elicited by the pictures were from the same set as those elicited by the digits. This latter finding is very important because the suppression mechanism is said to operate over nontarget lexical representations that presumably compete for selection, and there is no apparent way to stipulate how the mechanism could be operational when naming some stimuli but not others. If successful selection of an L2 lexical node is achieved through the suppression of its L1 competitors, then we should have observed evidence of this suppression persisting onto the subsequent L1 naming trial regardless of whether the stimulus eliciting a response in L1 was an Arabic numeral or a pattern of dots. Hence, our first conclusion in this article is that the language suppression hypothesis is wrong at some fundamental level because it incorrectly predicts effects of suppression where none are found.

Do we need to posit language suppression to account for the asymmetrical switch cost obtained with bivalent stimuli? Here the short answer is apparently not. In Experiment 3, we contrasted switching performance for fast and slow words to see whether the continuous variable speed of response availability, which is unavoidably conflated with language membership in language switching experiments, might be sufficient to produce an asymmetrical switch cost. We found that it was. In a unilingual experiment, participants took longer to switch from color naming to

word naming if the target word-naming stimulus was a fast-response word. This pattern mirrors the pattern observed in the language switching paradigm and, it is important to note, cannot be attributed to suppression in any straightforward way (see *Discussion* in Experiment 3). Although at present we are uncertain how, exactly, this correlated variable contributes to the asymmetrical switch cost, the crucial finding is that manipulating continuous lexical variables that are unavailable to a suppression mechanism is sufficient to produce an asymmetrical switch cost with bivalent stimuli. Thus, it is not necessary to posit a lexical suppression mechanism to account for the asymmetrical switch cost. Hence, our second conclusion in this article is that the asymmetrical language switch cost, though consistent with the predictions of the language suppression hypothesis, is a stimulus-specific effect that masquerades as an effect of language suppression.

### *Accounting for the Asymmetrical Switch Cost Without Lexical Suppression*

How might one account for the asymmetrical switch costs obtained with bivalent stimuli without appealing to suppression of lexical representations? Although it is beyond the scope of this article to formulate a formal explanation of these effects, it seems clear (from Experiment 3) that the speed (or ease) with which a response becomes available for production may be a critical factor. Because speed of response availability constitutes a set of continuous variables (e.g., frequency, length, number of senses), it is not clear how suppression could operate over the fast words independently of the slow words; yet the crossover interaction observed in Experiment 3 suggests that the fast words were suppressed independently of the slow words. Clearly, though, suppression here is of a very different sort than that proposed by Green (1998) and others because there is no evidence in Experiments 1 and 2 that lexical suppression ever occurred. For this reason, the term *response blocking* as opposed to *suppression* may be more appropriate in that it indicates a process that affects specific responses at a stage well after lexical selection has taken place. What kind of mechanism could block responses and, it is important to ask, block fast responses independently of slow responses?

In recent articles (Finkbeiner and Caramazza, in press; Finkbeiner et al., in press; Miozzo & Caramazza, 2003), a response selection account of performance in the Stroop-like picture-word naming task has been proposed, and it may be possible that this proposal could be extended to account for the findings reported above. A central assumption of the response selection account is that in Stroop-like tasks where stimuli afford two possible responses, the speech production system automatically makes both responses available to an output (articulatory) buffer, thereby necessitating the rejection of one of those responses so that the target response may be articulated over the single output channel. In the case of the picture-word interference paradigm, the decision to reject the word-naming response, which is typically first in the queue, is modulated by its response relevance (Lupker, 1979) as well as by how quickly that response becomes available for production (Miozzo & Caramazza, 2003). In the case of the language switching task, a bivalent stimulus, such as a digit, serves to generate two equally relevant responses, one in L1 and one in L2, and so the decision to reject or articulate has to be made on the basis of the color of the naming cue (e.g., green equals English and

gray equals German). When the naming cues are consistent with the preceding trial, as they are on nonswitch trials, the response selection criteria are already established, and responses may be selected (or rejected) as quickly as they become available for production. In this case, L1 responses are selected more quickly than L2 responses because L1 responses become available for production more quickly. When naming cues suddenly change, though, as they do on switch trials, it may be that some time is necessary before individuals are able to determine which response selection–rejection criteria are appropriate for that trial. Critically, on trials in which the response selection criteria must be reestablished (i.e., switch trials), we suggest that responses may inadvertently be rejected simply by becoming available for production too soon. The intuition here is that the more quickly (easily) a response becomes available in the context of a difficult switch trial, the more likely it is that participants will be suspicious of that response and dismiss it before it is articulated so as to ensure that an error is not made.

This proposal is similar to one made recently by Balota, Law, and Zevin (2000). In their study, half of the participants were asked to name words as quickly as possible, while the other half were asked to produce the “regular” (nonlexical) pronunciation of words (e.g., *pint* so that it rhymes with *mint*) as quickly as possible. Some of the words were irregular, like *pint*, whereas others had only a regular pronunciation (e.g., *weed*). Balota et al. (2000) found a normal frequency effect for regular words (e.g., *weed*) in the word-naming task but a reverse frequency effect for those same words in the regularization task. Balota et al. suggested that the more familiar (fast) the stimuli were, the more likely it was that they triggered a “double checking” procedure, whereby participants ensured that the pronunciation had been generated via a pronunciation rule (sublexical route) and not via the lexical route. In other words, the more quickly responses became available, the more likely it was that participants would be suspicious of those responses in the context of a difficult task.

The findings reported in the present article are consistent with this response selection proposal. The asymmetrical switch cost obtained with bivalent stimuli in Experiments 1–3 follows naturally from the response selection account because, to the extent that fast (L1) responses typically become available before slow (L2) responses, they are more likely to be inadvertently rejected on trials in which the response selection criteria suddenly change. It is important to note that because a rejected response must be regenerated before it can be produced, a greater time cost will be associated (counterintuitively) with the responses that become available more quickly. In the case of the univalent stimuli used in Experiments 1 and 2, the stimuli have only one response associated with them during the course of the experiment; hence, participants should not need any time to determine the appropriate response selection criteria for these stimuli and should be able to produce responses equally fast regardless of the response language on the previous trial. This is what we found in Experiments 1 and 2.

The response selection proposal is admittedly a very speculative proposal of switching costs at this point and is not the only possible account of the findings reported here. Waszak, Hommel, and Allport (2003) have recently proposed an explanation of task switch costs that may also be able to account for these findings. Waszak et al. (2003) suggested that executing a goal-directed action in response to a particular stimulus serves to create a

representation or “binding” between that stimulus and its corresponding response and that this *stimulus response event binding* may then be retrieved from memory later on by a re-presentation of the stimulus. Because this stimulus–response event binding also encodes information about the to-be-ignored response, re-presenting a stimulus will produce interference if the previously not-to-be-executed action is presently the appropriate action. Insofar as the stimulus–response event binding elicited from memory by the univalent stimuli in Experiments 1 and 2 was not associated with competing to-be-ignored responses, this strikes us as a reasonable account of why the univalent stimuli did not exhibit an asymmetrical switch cost and the digits did. That is, one could imagine that the degree to which the not-to-be-executed action is encoded into the stimulus–response event binding, or memory trace, differs depending on whether the to-be-ignored response is a preferred or dominant response for the stimulus. On switch trials, the dominant action for a particular stimulus (i.e., response) may take longer to execute than the nondominant action because it is encoded more strongly into the memory trace as the not-to-be-executed action. This may also be true for word-naming responses in the color–word naming task (Experiment 3). Essentially, the faster a word-naming response is available for production normally, the more strongly that response becomes encoded into the color-naming event file as the not-to-be-executed action (though note the lack of an effect in Experiment 3 on the color-naming trials). Upon switching from color naming to word naming, then, it may take longer to perform the word-naming action for fast words because those stimuli have been encoded more strongly as not-to-be-named stimuli.

Although both the response selection proposal and the account proposed by Waszak et al. (2003) offer possible explanations of the asymmetrical switch costs reported in the present article, it is very important to note how neither account appeals to the lexical selection mechanism. As such, neither proposal has any bearing on the question motivating the research here: How do bilinguals solve the hard problem? Although it may be possible to formulate a satisfactory explanation of the asymmetrical switch cost obtained with bivalent stimuli by appealing to response blocking processes or stimulus–response events encoded into memory, it does not appear possible to account for the asymmetrical switch cost by appealing to a suppression mechanism that operates at the point of lexical selection. Experiments 1 and 2 make this point most clearly. Thus, all that we can conclude definitively at this point is that neither the language suppression hypothesis (Green, 1998) nor its derivative, the lexical suppression hypothesis, offers plausible solutions to the hard problem in bilingual lexical access.

### *Implications for Models of Bilingual Speech Production*

Having rejected the language suppression hypothesis as a possible solution to the hard problem, how else might one model bilingual lexical access? Finkbeiner et al. (in press) and La Heij (2005; see also Poulisse & Bongaerts, 1994) have chosen to address the hard problem in a slightly different way. These authors have suggested that differences in the activation levels of translation-equivalent nodes may be achieved by activating the target node more than the nontarget node. According to these authors, suppression is not necessary. This *differential activation* approach (as opposed to *differential inhibition*) is quite possibly

the oldest proposal in the bilingualism literature. For example, the *language switch* proposal (Macnamara, 1967; Penfield & Roberts, 1959) and the *continuously monitoring operating system* proposal (Albert & Obler, 1978; Obler & Albert, 1987) hold that the intentions of the speaker serve to directly activate one language instead of, or more strongly than, the other. Today we know that the language switch hypothesis is wrong in its strongest form: Language systems are not turned on and off. Finkbeiner et al. (in press) have argued, though, that a weaker version of the language switch hypothesis, where the intentions of the speaker serve to activate the target language more strongly than the nontarget language, is a viable option. This weaker version is also seen in recent proposals by La Heij (2005) and Poulisse and Bongaerts (1994), who have similarly argued that the target language is specified at the conceptual level and that this specification serves to activate lexical nodes in the target language more strongly than their equivalents in the nontarget language.

Finkbeiner et al. (in press) have gone one step further by suggesting that the hard problem, which different models of bilingual lexical access have been designed to solve (e.g., Costa et al., 1999; Green, 1998; La Heij, 2005), may not be very hard after all. They have questioned whether the main assumption that gives rise to the hard problem, namely, the assumption of lexical selection by competition, is a valid assumption. If this assumption is shown to be unnecessary (cf. Finkbeiner & Caramazza, in press), then it is possible that lexical selection proceeds on the basis of a simple threshold mechanism, such as the one proposed by Dell (1986) or Caramazza and Hillis (1990). If one assumes that lexical selection operates on the basis of a threshold mechanism, then all that is necessary to ensure appropriate language selection is to stipulate that the bilingual's intention to speak in a particular language is able to modulate the rate at which activation accrues over target-language lexical nodes. There is much work yet to be done on this possibility, but Finkbeiner et al. (in press) have suggested that this simple proposal of bilingual lexical selection be convincingly rejected before more complicated accounts of bilingual lexical access are developed.

### Conclusion

The hard problem in bilingual lexical access arises when two lexical representations, one in L1 and one in L2, are activated to roughly equal levels and, thus, compete equally for lexical selection. The language suppression hypothesis (Green, 1998) has been proposed as a solution to this hard problem. According to the language suppression hypothesis, the selection of a target lexical node is facilitated through the suppression of the nontarget language and that (in unbalanced bilinguals) more suppression is needed to inhibit L1 than is needed to inhibit L2. The asymmetrical language switch cost is the signature effect of language suppression. This effect is characterized by participants taking longer to switch to their L1 (the more suppressed language) than their L2 (Meuter & Allport, 1999). The central finding in this study is that L2 learners who exhibited an asymmetrical language switch cost in a digit-naming task did not exhibit a language switch cost when switching from digit naming in L2 to picture naming in L1. Participants named pictures (in their L1) equally fast regardless of whether the previous digit-naming trial had been named in the participant's L1 or L2. This was true even when the pictures and

the digits elicited the same verbal responses (Experiment 2). This finding constitutes a serious challenge to the language suppression hypothesis (Green, 1998) because this hypothesis stipulates that the selection of an L2 lexical representation depends on the suppression of its L1 competitors and that this suppression persists onto the beginning of the next trial. Hence, we conclude that the language suppression hypothesis is wrong at a fundamental level because it predicts effects of suppression where none are found; consequently, we suggest that the language suppression hypothesis represents an unlikely solution to the hard problem in bilingual speech production.

### References

- Albert, M., & Obler, L. (1978). *The bilingual brain*. New York: Academic Press.
- Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 421–452). Cambridge, MA: MIT Press.
- Allport, A., & Wylie, G. (1999). Task-switching: Positive and negative priming of task-set. In G. W. Humphreys, J. Duncan, & A. M. Treisman (Eds.), *Attention, space and action: Studies in cognitive neuroscience* (pp. 273–296). Oxford, England: Oxford University Press.
- Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX Lexical Database (Release 2) [Computer software]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.
- Balota, D., Law, M., & Zevin, J. (2000). The attentional control of lexical processing pathways: Reversing the word frequency effect. *Memory & Cognition*, 28, 1081–1089.
- Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14, 177–208.
- Caramazza, A., & Hillis, A. E. (1990). Where do semantic errors come from? *Cortex*, 26, 95–122.
- Costa, A., Miozzo, M., & Caramazza, A. (1999). Lexical selection in bilinguals: Do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language*, 41, 365–397.
- Costa, A., & Santesteban, M. (2004). Lexical access in bilingual speech production: Evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language*, 50, 491–511.
- de Groot, A. M. B., & Christoffels, I. (in press). Language control in bilinguals: Monolingual tasks and simultaneous interpreting. *Bilingualism: Language and Cognition*.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 3, 283–321.
- Fellbaum, C. (Ed.). (1998). *WordNet: An electronic lexical database*. Cambridge, MA: MIT Press.
- Finkbeiner, M. (2003). Bilingual lexical memory: Towards a psycholinguistic model of adult L2 lexical acquisition, representation, and processing (Doctoral dissertation, University of Arizona, 2002). *Dissertation Abstracts International*, 63, 6114B.
- Finkbeiner, M., & Caramazza, A. (in press). Now you see it, now you don't: On turning semantic interference into facilitation in a Stroop-like task. *Cortex*.
- Finkbeiner, M., Forster, K. I., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language*, 51, 1–22.
- Finkbeiner, M., Gollan, T., & Caramazza, A. (in press). Lexical access in bilingual speakers: What's the (hard) problem? *Bilingualism: Language and Cognition*.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35, 116–124.

- Green, D. W. (1986). Control, activation and resource. *Brain and Language*, 27, 210–223.
- Green, D. W. (1993). Towards a model of L2 comprehension and production. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 249–277). Amsterdam: John Benjamins.
- Green, D. W. (1998). Mental control of the bilingual lexicosemantic system. *Bilingualism: Language and Cognition*, 1, 67–81.
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33, 149–174.
- La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism* (pp. 289–307). Oxford, England: Oxford University Press.
- Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22, 1–38.
- Lupker, S. J. (1979). The semantic nature of response competition in the Picture–Word Interference Task. *Memory & Cognition*, 7, 485–495.
- Macnamara, J. (1967). The linguistic independence of bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 6, 729–736.
- Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming: Asymmetrical costs of language selection. *Journal of Memory and Language*, 40, 25–40.
- Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in the picture–word interference paradigm. *Journal of Experimental Psychology: General*, 132, 228–252.
- Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7, 134–140.
- Monsell, S., Yeung, N., & Azuma, R. (2000). Reconfiguration of task-set: Is it easier to switch to the weaker task? *Psychological Research*, 63, 250–264.
- Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 4, pp. 1–18). New York: Plenum Press.
- Obler, L., & Albert, M. (1987). A monitor system for bilingual language processing. In M. Paradis (Ed.), *Aspects of bilingualism* (pp. 156–164). Columbia, SC: Hornbeam Press.
- Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton, NJ: Princeton University Press.
- Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 2, 539–557.
- Potter, M. C., So, K.-F., Von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and more proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 23, 23–38.
- Poulisse, N. (1997). Language production in bilinguals. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 201–224). Mahwah, NJ: Erlbaum.
- Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, 15, 36–57.
- Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory and Language*, 46, 245–266.
- Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107–142.
- Rogers, R., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: General*, 124, 207–231.
- Spector, A., & Biedermann, J. (1976). Mental set and mental shift revisited. *American Journal of Psychology*, 89, 669–679.
- Tokowicz, N., Kroll, J. F., de Groot, A. M. B., & Van Hell, J. G. (2002). Number-of-translation norms for Dutch–English translation pairs: A new tool for examining language production. *Behavior Research Methods, Instruments, & Computers*, 34, 435–451.
- Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology*, 46, 361–413.
- Yeung, N., & Monsell, S. (2003a). The effects of recent practice on task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 919–936.
- Yeung, N., & Monsell, S. (2003b). Switching between tasks of unequal familiarity: The role of stimulus-attribute and response-set selection. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 455–469.

## Appendix

### Materials Used in Experiment 3

Experimental items		Filler items	
“Fast” items	“Slow” items		
Rich	Wealthy	Origin	Native
Large	Huge	Emission	Burden
House	Cottage	Device	Debate
Material	Fabric	Culture	Severe
Small	Tiny	Driver	Comfort
Stone	Pebble	Aspect	Safety
Dog	Puppy	Breath	Entrance
Cat	Kitten	Mixture	Occasion
Fly	Mosquito	Damage	Magazine
Fast	Rapidly	Filling	Distinct

Received August 15, 2005

Revision received March 15, 2006

Accepted March 21, 2006 ■