Bilingualism: Language and Cognition 9 (2), 2006, 153–166 © 2006 Cambridge University Press doi:10.1017/S1366728906002501

Lexical access in bilingual speakers: What's the (hard) problem?*

MATTHEW FINKBEINER
Harvard University
TAMAR H. GOLLAN
University of California, San Diego
ALFONSO CARAMAZZA
Harvard University

Models of bilingual speech production generally assume that translation equivalent lexical nodes share a common semantic representation. Though this type of architecture is highly desirable on both theoretical and empirical grounds, it could create difficulty at the point of lexical selection. If two translation equivalent lexical nodes are activated to roughly equal levels every time that their shared semantic representation becomes activated, the lexical selection mechanism should find it difficult to "decide" between the two (the "hard problem") – yet in some cases bilinguals benefit from the presence of a translation equivalent "competitor". In this article, we review three models that have been proposed as solutions to the hard problem. Each of these models has difficulty accounting for the full range of findings in the literature but we suggest that these shortcomings stem from their acceptance of the assumption that lexical selection is competitive. We argue that without this assumption each proposal is able to provide a full account of the empirical findings. We conclude by suggesting that the simplest of these proposals should be rejected before more complicated models are considered.

Introduction

When wanting to communicate a particular concept through speech, it is first necessary to retrieve the lexical item that corresponds to the target concept. This process is commonly referred to as LEXICAL ACCESS and the way in which this process proceeds is far from straightforward. For example, when trying to name a picture of a cat, the intended semantic representation CAT becomes active, but closely related ones, such as DOG, PURR, TAIL, FUR, also become active (Dell, 1986; Caramazza, 1997; Levelt, Roelofs and Meyer, 1999). In most accounts, activated semantic representations are thought to spread activation down to their corresponding lexical representations at the lexical level. Because of this assumption these models require a lexical selection mechanism to decide which of the activated lexical representations should be chosen for further processing. Although there is still widespread disagreement over the precise way in which this mechanism works, most models of word production have shared the assumption that lexical selection is a competitive process (Starreveld and La Heij, 1996; Levelt et al., 1999; but see Dell, 1986; Caramazza and Hillis, 1990). According to this "selection by competition" assumption, the ease with which a target lexical node is selected depends not only on its own activation level, but on the activation level of competing lexical nodes as well. This is captured in the so called Luce ratio which states that there is an inverse relationship between the time required to select a target lexical node and the relative activation levels of competing lexical nodes. That is, as the difference between the activation levels of target and non-target lexical nodes decreases, it becomes increasingly difficult (and time consuming) to select the target lexical node.

Following directly from this central assumption of how lexical selection proceeds is the prediction that the closer two lexical representations are in meaning the more difficult it will be to select the correct one. For example, when presented with a picture of a couch in a picture naming experiment, the lexical representations *couch* and *sofa* should become activated to roughly equal levels because of their synonymy (Peterson and Savoy, 1998) and, hence, it should be very difficult to select the target lexical node. We refer to this difficult computational problem as the "hard problem". In bilingual speakers, the hard problem is extensive because virtually every concept is associated with synonymous lexical nodes. That is, many concepts, especially concrete concepts (Tokowicz, Kroll, De Groot and Van Hell, 2002), are associated with two translation equivalent lexical nodes in the bilingual mind. For this reason, the hard problem is one of the central questions in research on bilingualism.

It should be noted that the hard problem is only "hard" from a modeling perspective. Proficient bilinguals, for

* The preparation of this article was supported in part by a Career Development Award from NIDCD DC00191 to Tamar H. Gollan and by NIH grant DC 04542-04 to Alfonso Caramazza. We thank Albert Costa, Ton Dijkstra and two anonymous reviewers for their comments on an earlier version of the manuscript.

Address for correspondence
Matthew Finkbeiner, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA
E-mail: *msf@wjh.harvard.edu*

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press

whom lexical selection should be most difficult,[1] do not find it difficult to speak in one language instead of the other. Hence, either bilinguals have somehow overcome the hard problem or the assumption that lexical selection is competitive may be wrong and in need of revisiting. Below we review and discuss three different proposed solutions to the hard problem. All three proposals share as their starting point the assumption that lexical selection is competitive. With this assumption in mind, there are two logical solutions to the hard problem: either ensure that the lexical selection mechanism only considers the activation levels of lexical nodes in the target language, or ensure that the activation levels of lexical nodes in the target and non-target languages are prevented from approximating each other. In the first account that we review, Costa, Caramazza and colleagues proposed that bilinguals overcome the hard problem by only considering lexical nodes in the target language for selection. According to this proposal, the computational difficulty of the hard problem is never encountered because, though lexical selection is competitive within languages, it is not competitive across languages (Costa and Caramazza, 1999; Costa, Miozzo and Caramazza, 1999). In the second account that we will review, La Heij (2005) suggests that the activation levels of translation equivalent lexical nodes never come to approximate each other (the hard problem) because the semantic system activates target-language lexical nodes to a substantially higher level than lexical nodes in the non-target language. In the third proposal, which was developed over the years by Green (1986, 1993, 1998), the hard problem is solved by lowering the activation (through suppression) of the lexical nodes in the non-target language.

As our review of these three proposals will show, each one has been successful in accounting for particular phenomena, but none of the proposals has been successful in providing a complete account of the theoretical and empirical issues at hand. We will suggest that the limitations of each proposal stem not from their proposed solution to the hard problem, but from their acceptance of the selection by competition assumption. Abandoning this assumption, we will argue, allows each proposal to provide a full account of both the empirical and theoretical issues. In fact, we will propose that the three solutions to the hard problem are functionally indistinguishable from one another once the selection by competition assumption is dropped and that the thrust of future work should be directed towards finding new empirically tractable ways to adjudicate between the three proposals.

[1] The hard problem arises to the extent that the semantic system activates translation equivalent lexical nodes equally; and this is most likely to happen with proficient bilinguals because, for these bilinguals, the connections from the semantic system to lexical system are thought to be equally strong for L1 and L2 (Potter, So, Von Eckardt, and Feldman, 1984; Kroll and Stewart, 1994).

Language-specific lexical selection

The first proposal we consider was proposed by Costa, Caramazza and colleagues (Costa et al., 1999; Costa and Caramazza, 1999) to explain a striking set of results in the bilingual version of the picture word interference paradigm (Lupker, 1979; Glaser and Glaser, 1989). In this task, participants are asked to name a picture and to ignore a superimposed distractor word. A well-established finding with monolingual participants is that it takes longer to name a picture (e.g. dog) when the superimposed word is categorically related to the picture (e.g. "cat") relative to an unrelated baseline condition (e.g. "mat"). This interference effect has been taken as *prima facie* evidence that lexical selection is competitive. The reasoning is as follows: in the related condition, the lexical representation corresponding to the distractor word (*cat*) is thought to receive activation from two sources – from the written word and from the picture (via spreading activation at the semantic level); in the unrelated condition, the lexical representation corresponding to the distractor word (*mat*) only receives activation from the written word. Hence, the related context word "cat" will be a stronger competitor than "mat" because the activation level of its lexical representation (*cat*) will be higher than that of *mat*.

Because translation equivalents (e.g. *dog* and the Spanish equivalent *perro*) share a common semantic representation, they are even more closely related than semantically related words within a single language (e.g. *dog* and *cat*). As such, the selection by competition account predicts that the picture–word interference effect should be greatest when pictures are displayed with translation equivalent distractors. For example, when Catalan-Spanish bilinguals name a picture of a table in Catalan (*taula*), the Spanish translation equivalent (*mesa*) should slow naming responses significantly more than semantically related distractors (in either language). Impressively, however, Costa, Caramazza and colleagues found that the interference effect reverses completely and translation equivalent distractors lead to FASTER picture naming times relative to unrelated distractors (Costa and Caramazza, 1999; Costa et al., 1999).[2] These findings challenged the assumption that lexical selection is competitive across languages and led Costa, Caramazza and colleagues to propose that the lexical selection mechanism must not consider the activation levels of lexical representations belonging to the non-target

[2] Similar translation facilitation effects were found when translations were not presented explicitly. Bilinguals named pictures that they either could or could not translate (and that were matched for difficulty using monolingual data). Bilinguals named pictures more quickly (Gollan, Montoya, Fennema-Notestine and Morris, 2005) and with fewer TOTs (Gollan and Acenas, 2004) when they knew the translations. This suggests that translations facilitate (and do not inhibit) retrieval.

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press

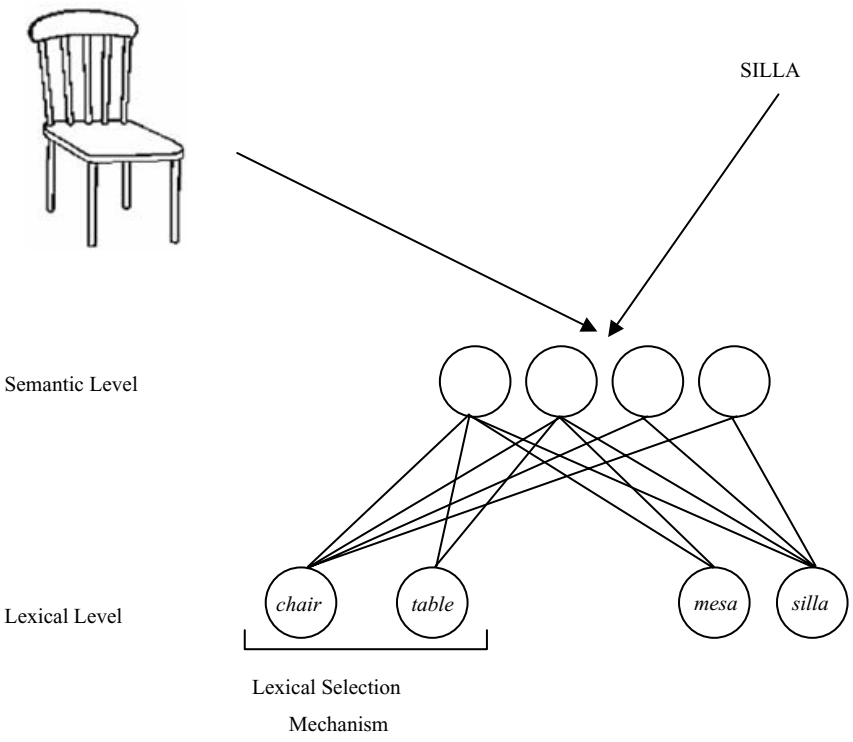


Figure 1. The language-specific lexical selection mechanism (Costa et al., 1999). According to this account, the lexical selection mechanism only considers for selection lexical representations belonging to the target language. In the picture–word interference paradigm, distractor words are thought to activate semantic representations directly (via an input orthographic lexicon).

language (see Figure 1). To account for the facilitatory translation identity effect, they proposed that, without competition across languages at the lexical level, a translation distractor word essentially acts as an identity prime at the semantic level. That is, the distractor word ("mesa") contributes to the activation level of its translation equivalent lexical node *taula* through its activation (priming) of the shared conceptual node TABLE.

In addition to the facilitatory translation identity effect obtained in the picture–word naming paradigm, Costa and Caramazza (1999) also observed a robust semantic interference effect for semantically related distractors. For example, when naming a picture of a table in Catalan, semantically related distractors such as "silla" (Spanish for "chair") and "cadira" (Catalan for "chair") produced an equal amount of interference. Costa and Caramazza proposed that the cross-language and the within-language semantic interference effect is essentially the same effect. According to their proposal, distractor words activate lexical-semantic representations directly (see Figure 1), which means that the distractor words "silla" and "cadira" are equally capable of activating the lexical node *cadira*. And because lexical selection is competitive WITHIN languages, the non-target-language distractor word "silla" will produce the same amount of interference as the target-language distractor word "cadira" when naming a picture of a table in Catalan.

Although the language-specific lexical selection mechanism provides a simple and effective solution to the hard problem, there are concerns with their proposal. For example, there is an effect in the findings reported by Costa and Caramazza (1999) that does not appear to be consistent with their proposal. We will refer to this as the "language effect". The nature of this effect is that participants take longer to name pictures (in the unrelated condition) with target-language distractors than they do with non-target-language distractors (Costa and Caramazza, 1999). In their study, Costa and Caramazza reported, on average, that participants named pictures 45 ms slower when unrelated distractor words were presented in Spanish (the target language) versus English.[3] The language effect was found regardless of

[3] It should be noted that the language effect was much smaller in the experiments with Spanish–Catalan bilinguals (Costa et al., 1999). This may be because of the close similarity between Catalan and Spanish words, or because of proficiency differences between the Spanish–Catalan and Spanish–English bilinguals.

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press

whether Spanish was the participants' dominant or non-dominant language. Why is this effect problematic for the language-specific lexical selection proposal made by Costa and Caramazza? If a distractor word is capable of activating its corresponding lexical node and its translation-equivalent lexical node equally, target- and non-target-language distractors should produce an equal amount of interference. But this prediction is not supported by the findings.

Costa and Caramazza acknowledged the language effect and suggested it had to do with the Spanish distractor words being longer than their English equivalents. However, reconsidering this account, it seems unlikely that length alone could have accounted for such a large difference. To investigate this possibility, we recently re-analyzed naming latencies in the unrelated condition from several picture–word interference experiments with monolingual English speakers and found in a simultaneous regression analysis that length (range: 3–10 letters) accounted for only 0.01% of the variance when length and frequency (of the target and of the distractor word) were entered as independent variables. Furthermore, the language effect observed by Costa and Caramazza is unlikely to be attributed to the frequency of the distractor words because the same language effect was found regardless of whether the distractors were presented in the participants' dominant (frequent) or non-dominant (less frequent) language.

How then might the language-specific lexical selection model be modified to account for the language effect? One straightforward possibility is to have written distractors activate their corresponding lexical nodes directly (without semantic mediation) such that translation-equivalent lexical nodes are activated to a lesser degree than same-language lexical nodes. This modification allows for an account of the language effect because target-language distractor words will produce more activation in the target lexicon than non-target-language distractors. Note, though, that this particular implementation predicts an interaction whereby the semantic interference effect obtained with target-language distractors should be larger than that obtained with non-target-language distractors. That is, the distractor word "cadira" should have produced more interference than the word "silla" when naming a picture of a table in Catalan. But this is at odds with the empirical findings. Presently it is not clear if it is possible to implement the language effect into the model proposed by Costa and Caramazza WITHOUT implementing a difference in the magnitude of the semantic interference effect for target- and non-target-language distractor words. As we will see below the concept selection proposal has difficulty grappling with similar findings in the picture–word naming paradigm, especially the translation identity effect.

Concept selection

La Heij (2005) characterized the hard problem as not being that hard after all. He begins by challenging the assumption that activation flows freely to multiple lexical representations, which he refers to as the "easy access, complex selection" approach. Instead, he proposes that only selected concepts, or, more appropriately, preverbal messages, activate their corresponding lexical nodes. La Heij's approach solves the hard problem insofar as the lexical node corresponding to the intended message receives strong activation while unintended near synonyms do not receive much activation at all. La Heij (2005) refers to this as the "complex access, easy selection" approach (see also Poulisse and Bongaerts, 1994; Poulisse, 1997). As his terminology suggests, the locus of the hard problem has been shifted up to the level of semantics. The claim is that whenever speakers need to choose between seemingly equivalent words to express a concept, subtle differences in meaning do the work. For example, the message MOVE TOWARD ME can be expressed politely as in "Could you please come over here?" or rudely as in "Get over here!" When the speaker's intention is to express the preverbal message politely, different words are activated and subsequently selected compared to when the speaker's intention is to express the preverbal message urgently and rudely. La Heij (2005) argues that something similar is true in bilinguals. For example, when a bilingual intends to say something in French the preverbal message, by virtue of specifying the intended language, ends up activating French more strongly than English. In some ways this approach is straightforward. After all, the decision to communicate in one language over another is based upon conceptual variables that include, for example, knowing who the interlocutor is, what the preferred language of communication is, what the communicative context is, and so on. Hence, the preverbal message must include the intended language in some way.

Nevertheless, few researchers have been willing to specify language membership at the level of meaning because of the considerable evidence suggesting that meanings do not change by virtue of the language that is used to express them (Berlin and Kay, 1969; Heider, 1972; Finkbeiner, Nicol, Nakamura and Greth, 2002; Li and Gleitman, 2002; but see Boroditsky, 2001). La Heij (2005) was careful to avoid criticism on this front, though, by proposing that the preverbal message is compositional in nature. Included in the preverbal message is, at minimum, a language-non-specific lexical-semantic representation, which is shared between translation-equivalent lexical nodes, a conceptual feature specifying the intended register, such as formal or informal, and a conceptual feature specifying the intended language, such as Spanish or English. Hence, if the intention is to say "dog",

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press

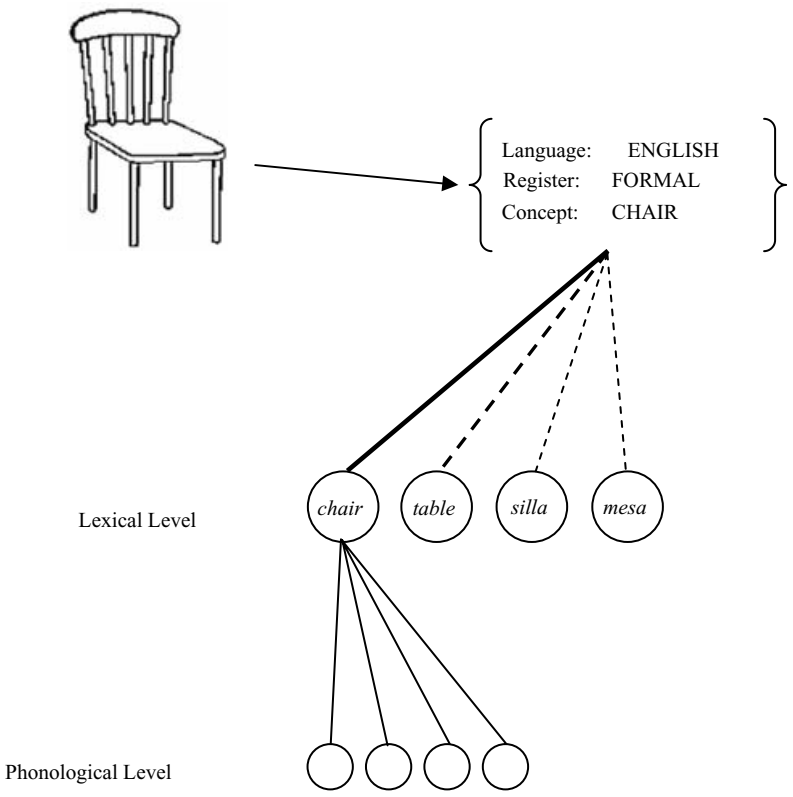


Figure 2. The "complex access, easy selection" proposal (La Heij, 2005). According to this proposal, the preverbal message specifies the lexical concept and the intended language, thus ensuring that the target lexical node receives the greatest amount of activation.

the preverbal message, by virtue of specifying English as the intended language, will ensure that the lexical node *dog* receives more activation than the lexical node *perro*. Following from this, lexical selection is a rather straightforward affair (see Figure 2). It should be noted that this proposal, like the one by Costa, Caramazza and colleagues, assumes that "selection" is language specific: the only difference is whether the selection has its effect at the conceptual (La Heij) or lexical level (Costa et al., 1999; Costa and Caramazza, 1999).

Just as there were concerns with the language-specific lexical selection mechanism proposed by Costa, Caramazza and colleagues, so are there concerns with the concept selection proposal suggested by La Heij. The first of these is more theoretical in nature. As pointed out above, one of the best-known and accepted findings in the lexical access literature is the picture–word interference (PWI) effect. Typically, as we discussed above, the PWI effect has been attributed to the lexical selection mechanism by saying that (a) lexical selection is a competitive process and (b) semantically related distractors are more potent competitors in the selection process because they receive activation from the target picture via the semantic system whereas unrelated distractors do not. In his new proposal, La Heij concurs that the locus of the PWI effect is at the lexical level. But, because he proposed that only selected concepts activate lexical nodes, he was forced to add the further stipulation that the preverbal message then activates a cohort of semantically related lexical nodes. That is, if the speaker's intention is to say "dog", the preverbal message (e.g. {concept: DOG; register: FORMAL; language: ENGLISH}) activates the lexical node *dog* most strongly, but it also activates the related lexical nodes *pooch*, *cat*, *fox* and *perro* among others. Hence, it is not clear how this differs from the more traditional model in which activation at the semantic level cascades down to the lexical level. Furthermore, it is difficult to imagine what would motivate the connection between the semantic node DOG and the lexical node *cat* (unless one assumed a distributed representation, which would then lead to the activation of all lexical nodes that are associated with any part of that representation).

The second and much more problematic concern for the concept selection proposal is the direction of the translation identity effect. Because, according to this proposal, the lexical selection mechanism considers the activation levels of lexical nodes in both lexicons

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press

simultaneously, translation identity distractors should induce much more interference than semantically related distractors. Yet, translation identity distractors FACILITATE picture naming processes. As such, it is unclear presently how the concept selection model could be modified to have its predictions square with the translation identity effect.

Distributed versus localist semantic representations

Although the concept selection model appears to have difficulty in accounting for the full range of empirical findings in bilingual speech production, a very attractive feature of La Heij's proposal is the role that the complex preverbal message plays in arriving at the correct lexical node. Intuitively, it seems very likely that conceptual information, including knowing who the interlocutor is, what languages she can speak, what register is appropriate when conversing with her, and so on, must serve to constrain the lexical representations that are entertained for selection. As we have seen, though, the concept selection model, as currently specified, is forced into the awkward position of having a single semantic node activate several related lexical nodes. How might a "complex access, easy selection" model avoid this awkward position? In the conclusion of his paper, La Heij (2005) suggested distributed semantic representations as opposed to localist representations as one possibility. Here we pursue this possibility briefly. If the preverbal message contained all of the necessary conceptual information to arrive at the correct lexical node, and was comprised of a bundle of conceptual features (Bierwisch and Schreuder, 1992; De Groot, 1993; Caramazza, 1997; Finkbeiner, Forster, Nicol and Nakamura, 2004), then any lexical node sharing any one or more of those conceptual features would become activated to some degree – the extent of which would depend upon the number of features in common. Following from this, the lexical nodes *cat* and *bark* and *wolf* will all become activated to differing degrees when the preverbal message includes the concept DOG. Furthermore, by assuming distributed semantic representations and, hence, one-to-many connections between conceptual features and lexical representations, the activation of these related lexical nodes is well motivated.

Although a model with distributed semantic representations may be able to account more straightforwardly for the activation of multiple lexical nodes by a single preverbal message, it is not any better off in explaining the facilitatory translation identity effects. In fact, any model that rests upon the selection by competition assumption will have difficulty accounting for this effect (without additional assumptions) unless it specifies language-specific lexical selection. And even then, as our review of Costa and Caramazza's model revealed, there are concerns with that particular implementation too. Below we review Green's Inhibitory Control Model, which proposes that lexical nodes in the non-target language are suppressed so that lexical selection may proceed in the target language without interference.

Selection through inhibition

According to Green's Inhibitory Control Model (ICM; 1986, 1993, 1998), each lexical representation is associated with a language tag, and lexical nodes can be suppressed if they are associated with the non-target language in a particular communicative context (see Figure 3). In this way, lexical selection proceeds without facing the hard problem because translation-equivalent lexical nodes, which would otherwise be just as highly activated as the target nodes, are suppressed. Three important features of the ICM are that (a) the semantic system activates lexical nodes in both languages (b) lexical nodes in the non-target language are then suppressed reactively, and (c) inhibition is proportional to the level of activation of the lexical nodes in the wrong (non-target) language – the more the semantic system activates representations in the wrong language, the stronger the inhibition of this language. According to the ICM, the semantic system activates L1 more strongly than L2 and because suppression is proportional to activation levels, L1 is thought to be the most strongly inhibited when it is not the target language. The L2, on the other hand, is not as highly activated and so it receives less inhibition when it is not the target language.

Green's Inhibitory Control Model (ICM) is unique in the bilingual speech production literature in that the bulk of support for this model comes from studies on cued language switching instead of picture naming, and, in particular, from the seminal study by Meuter and Allport (1999). In this study, participants were asked to name single digits (range: 1–9) presented on a screen in either their dominant language (L1) or non-dominant language (L2), depending upon the color of the screen. Meuter and Allport found that participants named digits faster in L1 than in L2 on trials in which no language switch was required, but (paradoxically) that participants took longer to switch to L1 than they did to switch to L2. Though these findings were viewed as somewhat counterintuitive, they followed directly from the predictions of the ICM. The reasoning is as follows. If L1 must be strongly inhibited to name a digit in L2 on trial N, then it should take some time to produce a response in L1 on the subsequent trial (N+1) because of the time needed to reactivate L1. Hence a large switch cost for L1. The L2 exhibits a much smaller switch cost in comparison because on trial N, when L1 is the response language, L2 is only minimally inhibited. Switching to L2 on trial N+1 and reactivating the L2 system is easily done (relatively

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press

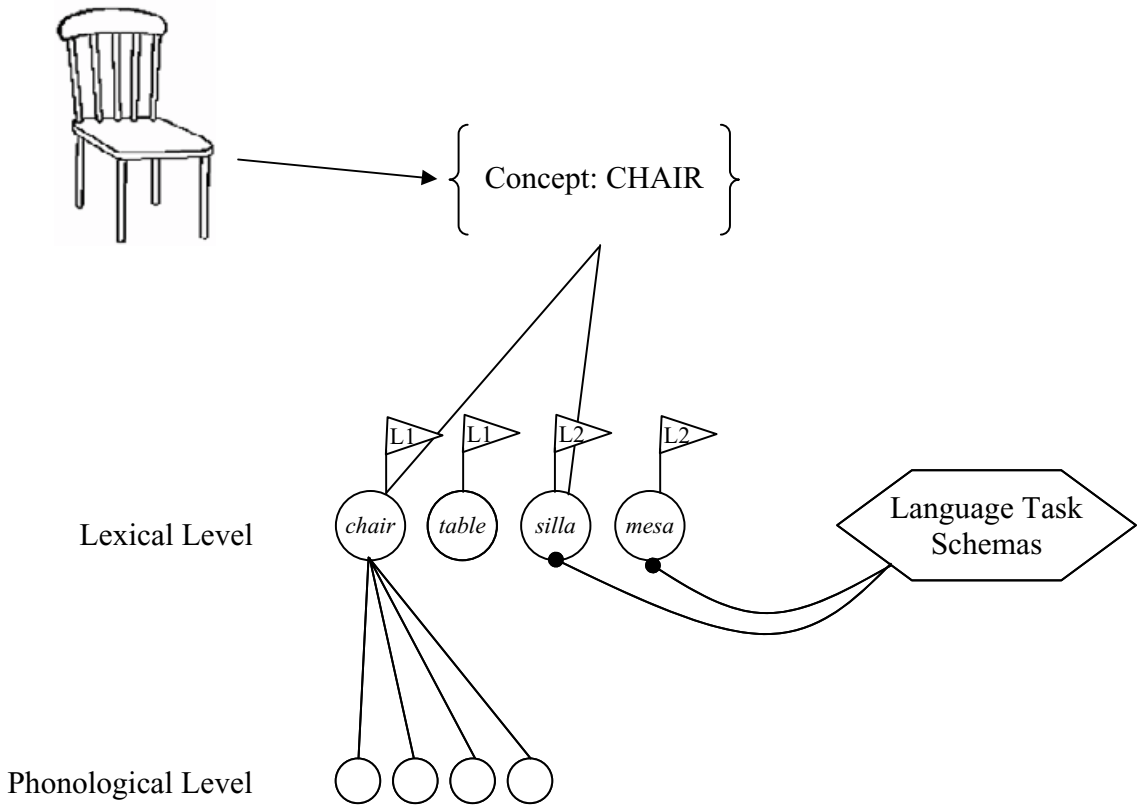


Figure 3. A simplified version of the Inhibitory Control Model (Green, 1998). In this example, the target language is English (the speaker's L1), and the non-target language is Spanish (L2). The inhibitory connections between the language task schemas and L2 lexical nodes indicate suppression of lexical nodes in the non-target language.

speaking) because it was only weakly inhibited in the first place.

The asymmetrical language switch cost reported by Meuter and Allport (1999) is quite robust and has been replicated several times in digit naming experiments. Nevertheless, some recent findings indicate that the language suppression hypothesis actually makes the wrong predictions in the language switching task. For example, Finkbeiner, Almeida, Janssen and Caramazza (in press) had L2 learners perform the same language switching task used by Meuter and Allport with the only difference being that digits and pictures (line drawings in one experiment and patterns of dots in another) were used to elicit responses. The pictures were always named in L1 while the digits were named in L1 and L2. Participants exhibited an asymmetrical language switch cost on the trials in which digits were used to elicit a response, effectively replicating the original findings reported by Meuter and Allport (1999), but did not exhibit a switch cost on the picture-naming trials. That is, participants named line drawings and dot patterns in their L1 equally fast regardless of whether the preceding digit-naming trial had been named in L1 or L2. This finding is highly problematic for the language suppression hypothesis because, according to this hypothesis, the suppression of an L1 lexical node should persist from the preceding L2 naming trial onto the subsequent L1 naming trial regardless of whether the stimulus used to elicit the L1 response is a digit or a picture. Apparently, though, the asymmetrical language switch cost, which is readily observed with stimuli that elicit both L1 and L2 responses (i.e. digits), does not extend to stimuli that elicit L1 responses only.

In another recent study, Costa and Santesteban (2004) reported two main findings that when taken together also constitute a serious challenge to the language suppression hypothesis proposed by Green (1998) and Meuter and Allport (1999). The first finding is that highly proficient bilinguals exhibited SYMMETRICAL switch costs. This was true even when these bilinguals were asked to switch between their L1 and their much less proficient L3. The second finding is that in each case of symmetrical switching costs, L1 responses were produced more slowly than L2 (L3) responses. This pattern of findings is problematic for the language suppression hypothesis because if one were to appeal to suppression to explain the slower L1 responses, one should predict a larger L1 switch cost – but switch costs were symmetrical for these bilinguals. Apparently, then, proficiency is an important variable in determining whether or not one

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press

obtains an asymmetrical switch cost in this task. Costa and Santesteban (2004) have suggested that highly proficient bilinguals, but not L2 learners, are able to modulate the availability of lexical items for lexical selection in each language separately. Although Costa and Santesteban's proposal will need to be tested further, their findings along with those reported by Finkbeiner et al. (in press) make clear that the suppression hypothesis makes the wrong predictions not only for L2 learners (see above) but for advanced bilinguals too.

Interim summary

Bilinguals confront the hard problem to a greater extent than monolinguals because in bilinguals all concepts are associated with synonymous and, thus, equally activated, lexical nodes. As such, extant models of lexical access predict that lexical selection in the bilingual case should be very difficult. Yet bilinguals clearly are not confronted with the hard problem when they speak because, for proficient bilinguals at least, speaking in one language and not the other is done easily and effortlessly. Clearly, then, bilinguals have solved the hard problem. Above we reviewed three models which represent three different classes of solutions to the hard problem. The first of these proposes that the lexical selection mechanism only considers one lexicon, the target lexicon, when selecting lexical nodes. The second model proposes that the target lexicon receives more activation from the semantic system than the non-target lexicon. And the third model proposed that lexical nodes in the non-target lexicon are suppressed.

As we have seen, each model has difficulty in accounting for the full range of relevant empirical findings. Below we argue that despite these apparent shortcomings, each class of models remains equally viable as a possible solution to the hard problem. We suggest that the shortcomings of the particular models we have reviewed stem from how the empirical findings marshaled in support of the individual models have been interpreted. It is important to point out that we do not dispute the empirical findings supporting the various models, such as the picture–word interference effect or the asymmetrical language switching cost, but we question whether those findings mean what they have been purported to mean (both with respect to constraining theories of bilingual lexical access and more generally with respect to how lexical selection proceeds in monolinguals). More specifically, we will argue that findings from Stroop-like paradigms, such as the picture–word naming paradigm and the language switching paradigm, do not provide adequate support for the selection-by-competition and language-suppression assumptions respectively. As such, we will argue that lexical selection processes may be reformulated to proceed via a simple threshold mechanism and that the hard problem may be solved simply by creating a difference in the rate (or level) of activation of target- and non-target-language lexical nodes or by selecting from only the target lexicon.

Reinterpreting the data: Do we really need selection by competition?

In a series of recent papers (Costa, Alario and Caramazza, 2005; Finkbeiner and Caramazza, in press; Finkbeiner et al., in press), Caramazza and his colleagues have developed a response selection account of performance in Stroop-like tasks as an alternative to the lexical selection by competition account (e.g. Levelt et al., 1999). We should be clear from the beginning of our discussion of the response selection account that it is not intended as a model of lexical access; rather, it is an attempt to explain how individuals perform Stroop-like tasks. Evidence from Stroop-like tasks has been used to motivate and constrain different models of lexical access, such as the three models reviewed above, and we question the validity of using unnatural interference-inducing tasks to inform a theory of normal lexical access. The motivation for developing the response selection account, then, is to determine whether it is possible to explain Stroop-like interference in terms of peripheral decision-level processes that, importantly, do not implicate the lexical selection mechanism. To the extent that this endeavor is successful, we may call into question the motivation for incorporating into the lexical selection mechanism the complicated assumptions and procedures that are necessary to account for performance in Stroop-like tasks. That is, if there is no need to appeal to the lexical selection mechanism to account for Stroop-like interference, then it seems reasonable to offer in its place a simpler mechanism. In the remainder of this article, we argue that performance in interference-inducing tasks, such as the picture–word interference task or the language switching task, may be accounted for straightforwardly by the response selection account. As such, we suggest that there is no compelling reason to retain the assumption of selection by competition and offer in its place a simpler "selection by activation threshold assumption". We argue that adopting the selection by activation threshold assumption has the consequence of making the hard problem in bilingual speech production relatively easy to solve.

The response selection account is based upon the following assumptions. First, in a Stroop-like naming task in which each stimulus affords two possible responses, we assume that individuals cannot help but to formulate (covertly) both phonological responses afforded by the stimulus. This is no different from the long-held assumption that subjects are unable to avoid reading the word distractors in the Stroop task. The second assumption follows from the limitation imposed by the architecture

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press

of the speech production system, which is that there is only one output channel and that only one response may be produced at a time over this channel. In the case of Stroop-like tasks where an individual cannot help but to generate two responses for a single stimulus, we assume that individuals must "decide" between responses and remove the inappropriate response from the output buffer so that the appropriate response may be produced over the output channel. There are three critical aspects to this response selection process. First, we assume that response selection processes operate over phonologically well-formed responses in an output buffer, not lexical nodes at a more abstract level of representation. Second, it is assumed that the decision mechanism operates over available responses in a serial fashion, beginning with the response that becomes available first. In the case of picture–word stimuli, this would be the response engendered by the distractor word. Third, following Lupker (1979), we assume that the speed with which the decision-level response selection process may be completed is modulated by whatever relevance the non-target response may have vis-à-vis the task at hand. Here, the assumption is that when a phonological response becomes available, it carries with it its categorical information, its provenance (word stimulus vs. picture stimulus), and its lexical information (e.g. grammatical class and language identity), all of which can be used by the response selection mechanism in its decision to accept or reject the response.

In Stroop-like tasks, a single mismatch between an available response and the response selection criteria established in a particular task (e.g. picture naming) can be used to reject a response. For example, if the task is to name pictures in English and the word-naming response (which becomes available before the picture-naming response) is Spanish, it can be rejected quickly and before it engages selection processes any further. In contrast, if the word-naming response meets several of the selection criteria, then it will take longer to reject and, hence, will delay production of the picture-naming response.

Several recent findings provide support for the response selection account. For example, in the picture–word interference (PWI) task, in which participants are asked to name a picture and to ignore a superimposed distractor word, Miozzo and Caramazza (2003) reported that low frequency distractor words interfere more than high frequency words in the PWI task. This finding, which is very robust, is quite counterintuitive if one assumes that the relative activation level of a lexical competitor modulates lexical selection times. This is because high frequency distractors, by virtue of having relatively high levels of activation, should produce more (or, at the very least, NOT LESS) interference than low frequency distractors. Nevertheless, this finding follows directly from the response selection account. Essentially, the sooner a distractor can be rejected from production, the sooner the picture name can be selected and because high frequency words become available for production sooner than low frequency words, the response selection mechanism is able to reject high frequency distractor words sooner.

Another line of support comes from recent findings which reveal that only a very specific type of semantic relationship between picture and distractor induces the picture–word interference effect. The selection by competition account predicts that all semantically related words should cause more interference than unrelated words because related words receive activation spreading from the picture and, hence, should be more highly activated – i.e. able to compete more strongly. However, Costa et al. (2005) found that although co-ordinate (car – truck) pairs lead to interference, "has-a" (car – bumper) picture–word pairs FACILITATED responses. Here we suggest that it takes longer to reject co-ordinate distractors because the target and distractor share categorical information and thus the decision mechanism cannot rely on mismatching categorical information alone to reject the distractor. In the case of "has-a" distractors, these words (e.g. bumper) should not compete because they have little relevance vis-à-vis the task of naming whole objects (e.g. car). Why, though, did "has-a" distractors FACILITATE picture naming? In the absence of any response relevance, the cost attributed to response selection processes is minimal and, crucially, is less than the benefit conferred by semantic priming. That is, a distractor word is processed at two distinct levels with two distinct consequences. At the semantic level, all activation is assumed to be facilitatory, and so semantic relatedness between distractor and target leads to priming at this level. At the level of response selection, shared category membership between distractor and target eliminates a feature that the selection mechanism could have otherwise used to reject the word-naming response; this has the consequence of delaying the rejection of that response.

In another line of work (Finkbeiner and Caramazza, in press), we found that the interference effect reverses completely if distractor words are presented very briefly and masked so that participants cannot report their presence. That is, the very same semantically related words that interfere with picture naming processes when they are readily visible actually FACILITATE naming latencies when they are masked. Again, the selection by competition assumption holds that semantically related words should COMPETE for selection more than unrelated words. Because we observed robust priming effects, we can conclude that the masked primes produced activation within the lexical system, and yet we observed facilitation, not interference, from the masked distractors.

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press

Here we suggest that the masking procedure prevents the formulation of phonologically well-formed responses, and, thus, prevents decision-level response-selection processes from becoming engaged.[4] This, in turn, leaves behind only the facilitatory effects of semantic priming. That is, the masking procedure effectively turns the bivalent, Stroop-like stimulus with two possible responses into a univalent stimulus, which has only one response (the picture naming response) associated with it.[5]

Of course, though the picture–word interference task has provided the bulk of support for the lexical selection by competition assumption, findings in other experimental paradigms have also been interpreted in terms of this assumption. For example, in the semantic blocking paradigm where participants are asked to name a series of pictures that are either taken from the same semantic category (homogenous condition) or different categories (heterogeneous condition), participants take longer to name pictures in the homogenous condition (Kroll and Stewart, 1994; Damian, Vigliocco and Levelt, 2001). These findings have several potential explanations, one of which is consistent with the response selection proposal described above. For example, it seems reasonable to consider the possibility that the response produced on trial N would still be available for production on trial N+1 and, hence, may enter into the response selection process on a proportion of trials. If this were the case, then it is not surprising that interference is observed in the homogenous condition because of the shared category membership for the two responses.

Turning now to the bilingual version of the picture–word interference task, there are three important data points that need to be explained, all of which find a ready explanation in the response selection account. The first of these is the main effect of language, which has not received much attention in the literature previously. The "response selection" proposal has a ready account for the language effect: namely, non-target-language distractors meet fewer response selection criteria than target-language distractors and, hence, are able to be rejected sooner in the decision process than target-language distractors. As such, participants are able to name pictures in the presence of non-target-language distractors more quickly than they can in the presence of target-language distractors.[6]

The second data point, which has received much more attention in the literature, is the semantic interference effect for related target- and non-target-language distractors. An important feature of this effect is that the magnitude is identical for both target- and non-target-language distractors. The response selection approach proposed here can account for this data point straightforwardly. Insofar as translation-equivalent distractors share a common referent, target- and non-target-language distractors will be equally similar (or dissimilar) to the referent of the picture. Hence, although non-target-language distractors should be discarded as possible responses faster than target-language distractors (see preceding paragraph), the relative difference in semantic similarity between related and unrelated distractors is the same for distractors in both languages. Thus, insofar as semantic similarity between picture and word slows the decision process to reject a word naming response, the relative effect of semantic similarity should be identical for translation-equivalent distractors.

The third data point is the identity effect for target- and non-target-language distractors. Typically, the within-language identity effect (e.g. name a picture of a dog in English with the word *dog* superimposed on the picture) is approximately 135 ms and the translation identity effect (e.g. name a picture of a dog in English with the word *perro* superimposed on the picture) is approximately 30 ms.[7] First, a proposal such as La Heij's or Green's, which assumes competition across languages, says that non-target-language identity distractors should compete with target selection, not facilitate it. Interestingly, the response selection proposal made here also suggests that non-target-language distractors should "interfere" with target naming responses, but our proposal, unlike those based upon the selection by competition assumption, predicts interference relative to the target-language identity condition only. According to the response selection proposal, naming is fastest with target-language identity distractors (e.g. DOG – *dog*) because of priming at the semantic level and because nothing needs to be rejected at the response level in this condition. The translation equivalent distractor condition (e.g. DOG – *perro*) is next slowest because, while a response needs to be rejected, it is available for rejection sooner than unrelated cross-language distractors because of the semantic priming produced by the picture. In the unrelated condition, the

[4] Note that this does not imply that we presume that the masking procedure is effective in preventing any activation at the phonological level; rather, we simply propose that the masking procedure is effective in preventing phonologically well-formed responses from becoming available to the response selection mechanism.

[5] According to the response selection proposal, the determiner congruency effect (Miozzo and Caramazza, 1999; Schiller and Caramazza, 2003), which has previously been cast in terms of competition between determiners in the gender-incongruent condition, will need to be reformulated. One possibility is that the congruency effect reflects priming in the gender-congruent condition as opposed to competition in the gender-incongruent condition.

[6] The language effect may also be explained by La Heij's concept selection model. According to La Heij's account, target-language distractors compete for selection more strongly because they are more highly activated than non-target-language distractors.

[7] The mean effect sizes referred to here were determined by averaging across the effect sizes reported in the relevant experiments from Costa and Caramazza (1999) and Costa, Miozzo and Caramazza (1999).

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press

non-target-language distractors should be next slowest. This condition is faster than unrelated target-language distractors for the same reason used to explain the language effect (first data point discussed above). The unrelated target-language distractors are slowest to be rejected because they do not benefit from priming and because they *do* constitute a valid response (by virtue of belonging to the target language). These predictions are confirmed by the empirical findings in the literature (e.g. Costa and Caramazza, 1999; Costa et al., 1999). Thus, the response selection proposal made here has a ready explanation for the facilitatory translation identity effect, the language effect, and the fact that the non-target-language identity effect is smaller than the target-language identity effect. As we indicated above, proposals that adhere to the selection by competition assumption either have difficulty accounting for the language effect (Costa and Caramazza, 1999), or they make the wrong predictions altogether for the translation identity condition (La Heij, 2005).

We point out that Hermans (2000) and Green (2002) have suggested that one may retain the lexical selection by competition assumption and still account for the facilitatory translation identity effect by appealing to semantic priming. These authors have both suggested that one way around the predicted interference effect for translation identity distractors is to stipulate that the cost associated with lexical selection by competition asymptotes quickly and at a value that is substantially smaller than the benefit conferred by the activation (priming) produced at the semantic level by the identity distractor. Though this is a possible solution to the facilitatory translation identity effect, it is unclear how this type of explanation could account for the language effect (data point #1) without predicting a difference (interaction) in the semantic effect for related target- and non-target-language distractors (data point #2).

We also point out that the so called "phono-translation" effect (Hermans, Bongaerts, De Bot and Schreuder, 1998; Costa, Colomé, Gómez and Sebastián-Gallés, 2003) may constitute a challenge to our account of the between-language semantic interference effects. In the phono-translation experiments, participants name pictures in one language with AUDITORILY PRESENTED distractor words that are phonologically related to the translation equivalent of the picture's name. For example, Costa et al. (2003) had Catalan–Spanish bilinguals name pictures in Catalan (e.g. *gos* for a picture of a dog) with Spanish distractors that were semantically related (*zorro* "fox"), phonologically related to the Spanish translation equivalent (*pelo* "hair" which is phonologically related to *perro*, Spanish for "dog"), phonologically related to the target (*gozo* "pleasure") and unrelated (*lluvia* "rain"). In the condition of interest, Costa et al. (2003) found that participants were reliably slower in the phono-translation condition (name *gos* in the presence of the distractor *pelo*) compared to the unrelated condition at the positive (+150 ms) SOA.[8] These findings effectively replicated those that were reported earlier by Hermans et al. (1998) with Dutch–English bilinguals. Hermans et al. interpreted their findings to mean that lexical selection was competitive and that the lexical selection mechanism was not language specific.

Presently we are not certain how to interpret these findings within the response selection framework. We suspect, though, that the phono-translation effect is due in large part to strategic effects (see footnote 8). For example, it may be that in the context of the picture, participants, on a number of trials at least, initially misperceive the phono-translation distractor as the name for the target (e.g. *pelo* as *perro*) but then realize their mistake, which incurs a time cost. One way to test this possibility would be to train participants extensively with the distractor words prior to the experiment proper. Our suspicion is that extensive training would have little effect on the magnitude of the semantic effect but would dramatically reduce the magnitude of the phono-translation effect.

In summary, we suggest that the response selection proposal is able to account for the naming performance in Stroop-like tasks without appealing to competitive lexical selection processes. As such, our reinterpretation of the empirical findings calls into question the relevance of participants' performance in Stroop-like tasks in constraining either monolingual or bilingual theories of lexical access. By extension, our reinterpretation of the findings used to support the lexical selection by competition assumption allows for a reformulation of lexical selection processes. In contrast to the selection by competition assumption, we suggest that lexical selection may proceed on the basis of a simple threshold mechanism (Morton, 1969; Dell, 1986; Caramazza and Hillis, 1990). Assuming this to be the case, then the solution to the "hard problem" is relatively easy: simply create a difference in the rate at which activation accrues over target- and non-target lexical nodes.

Putting it all together

The hard problem that models of bilingual speech production need to solve concerns how lexical selection occurs when the intention to speak activates translation-equivalent lexical representations to roughly equal levels. This has constituted a "hard problem" in the bilingualism literature because of the general acceptance of the assumption that lexical selection is competitive. That is,

[8] The phono-translation effect was reliable at all SOAs in Experiment 1 when no filler trials were included but it was only reliable at the positive SOA in Experiment 2 when filler trials were included.

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press

a fundamental assumption in the speech production literature is that lexical selection becomes more difficult as the difference between the activation levels of target and non-target lexical nodes decreases. Following from this, a proficient Spanish–English bilingual should find it very difficult to say "dog" because the intention to say "dog" activates the lexical node *perro* just as much as it does the lexical node *dog*. But this prediction is incorrect. Proficient bilinguals do not find it difficult to speak in one language only.

In this article, we reviewed three proposed solutions to the hard problem (Green, 1998; Costa and Caramazza, 1999; La Heij, 2005). Each of these three proposals is based upon the lexical selection by competition assumption. Costa and Caramazza (1999) address the hard problem by suggesting that the lexical selection mechanism does not consider the non-target lexicon. Although this proposal is able to account for the facilitatory translation identity effects in the bilingual version of the picture–word interference task, their model is not able to account for the "language effect": the finding that target-language distractors interfere more than non-target-language distractors. Also, it is not clear in their proposal HOW the selection mechanism knows which lexicon is the target lexicon. The concept selection model proposed by La Heij (2005) addresses the hard problem by suggesting that lexical nodes in the target lexicon are activated to a much higher degree than lexical nodes in the non-target lexicon. Although this model is intuitively very appealing, it is not clear within the constraints of the concept selection model how to account for the facilitatory translation identity effect. The Inhibitory Control Model proposed by Green (1998) addresses the hard problem by suggesting that lexical nodes in the non-target-language lexicon are suppressed. Although language suppression seems perfectly reasonable as a solution to the hard problem, it appears to be undermined by findings which reveal that language suppression is limited to just those instances in which bivalent stimuli are used (i.e. stimuli that elicit both L1 and L2 responses during the course of the experiment).

Given the difficulties that each of these models have in accounting for the full range of empirical findings, one possible course of action would be to modify the models so that they square with the empirical findings. A disadvantage of this approach, though, is that it would undoubtedly lead to an increase in the complexity of the models. Another possible approach, and one that we have advocated here, is to say that the empirical findings from dual-task paradigms such as picture–word interference and language switching do not constrain models of lexical access in the way that has been assumed in the literature. That is, those data do not speak directly to issues about bilingual lexical selection. We discussed an alternative "response selection" account of the effects observed in Stroop-like naming paradigms, including the picture–word interference task (cf. Miozzo and Caramazza, 2003; Costa et al., 2005; Finkbeiner and Caramazza, in press) and the language switching task (cf. Finkbeiner et al., in press), which implicates response selection processes, not lexical access processes.

What does shifting the locus of Stroop-like effects down to the response stage (output buffer) buy us in terms of a solution to the hard problem? Quite simply, it allows for an interpretation of these effects that does not implicate lexical access processes. Previously, for example, the picture–word interference task was taken as *prima facie* evidence that lexical selection was competitive (Roelofs, 1992; Starreveld and La Heij, 1996; Levelt et al., 1999). The reinterpretation of those data offered here suggests that participants' performance in Stroop-like tasks does not necessarily motivate the assumption that lexical selection is competitive. And without the motivation for lexical selection by competition, we propose that the lexical selection mechanism may be reformulated to operate on the basis of a simple threshold mechanism (Dell, 1986; Caramazza and Hillis, 1990). A SELECTION BY THRESHOLD mechanism simply stipulates that the first node whose activation level reaches threshold will be selected. Within this framework (Morton, 1969; Dell, 1986), the degree to which non-target lexical nodes are activated does not affect the time needed for the target node to accumulate enough activation to reach threshold. This reformulation of the lexical selection mechanism is extremely important. If lexical selection occurs on the basis of a threshold mechanism, then the solution to the hard problem is within reach. In fact, the hard problem should not be considered hard at all. According to the selection by threshold assumption, all that is needed to ensure correct selection of target-language lexical nodes is to increase the rate of activation of target-language lexical nodes relative to that of non-target-language nodes.

While selection of target-language lexical nodes could be ensured by only considering the target-language for selection (Costa and Caramazza, 1999) or by suppressing lexical nodes in the non-target language (Green, 1998), a simpler solution seems to be available: namely, let the intentions of the bilingual to speak in one language and not the other modulate the rate at which activation accrues over lexical nodes in the target and non-target language. We may refer to this as the "differential activation" proposal. This proposal is not new with us. In fact, it quite possibly is the oldest proposal. For example the "language switch" proposal (Penfield and Roberts, 1959; Macnamara, 1967) and the "continuously monitoring operating system" proposal (Albert and Obler, 1978; Obler and Albert, 1987) hold that the intentions of the speaker serve to direct activation to one language instead of, or more strongly than, the other. Today we know that the language switch hypothesis is wrong in its strongest form:

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press

language systems are not turned on and off. Nevertheless, it would seem 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 non-target language, is still a viable option. In fact, this very proposal has been incorporated into recent models of bilingual speech production. Poulisse and Bongaerts (1994) suggested 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 non-target language. And, as we have reviewed in this article, La Heij (2005) has similarly proposed that the speaker's intentions may serve to differentially activate the target language over the non-target language.

Conclusion

We reviewed three different proposed solutions to the hard problem in bilingual lexical access: the language-specific lexical selection mechanism (Costa and Caramazza, 1999), the inhibitory control model (Green, 1998) and the concept selection hypothesis (La Heij, 2005). As currently specified, each of these models has difficulty accounting for the full range of empirical findings from Stroop-like paradigms. We proposed a reinterpretation of the findings from these paradigms, arguing that these data are not necessarily informative with respect to the hard problem. If one accepts our reinterpretation of participants' performance in dual-task paradigms, then the three proposals that we reviewed all constitute logical solutions to the hard problem. But, if one accepts our reinterpretation of these data, one must also accept that all three proposals are presently without empirical support. We recommend that before considering more sophisticated proposals, future work on bilingual lexical access should be directed at rejecting the simplest possibility, which we call the "differential activation" hypothesis.

References

Albert, M. & Obler, L. (1978). *The bilingual brain*. New York: Academic Press.

Berlin, B. & Kay, P. (1969). *Basic color terms*. Berkeley, CA: University of California Press.

Bierwisch, M. & Schreuder, R. (1992). From concepts to lexical items. *Cognition*, 42, 23–60.

Boroditsky, L. (2001). Does language shape thought? English and Mandarin speakers' conceptions of time. *Cognitive Psychology*, 43 (1), 1–22.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive Neuropsychology*, 14, 177–208.

Caramazza, A. & Hillis, A. (1990). Where do semantic errors come from? *Cortex*, 26, 95–122.

Costa, A. & Caramazza, A. (1999). Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. *Bilingualism: Language and Cognition*, 2, 231–244.

Costa, A., Alario, F.-X. & Caramazza, A. (2005). On the categorical nature of the semantic interference effect in the picture–word interference paradigm. *Psychonomic Bulletin and Review*, 12, 125–131.

Costa, A., Colomé, A., Gómez, O. & Sebastián-Gallés, N. (2003). Another look at cross-language competition in bilingual speech production: Lexical and phonological factors. *Bilingualism: Language and Cognition*, 6, 167–179.

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.

Damian, M., Vigliocco, G. & Levelt, W. (2001). Effects of semantic context in the naming of pictures and words. *Cognition*, 81, B77–B86.

De Groot, A. M. B. (1993). Word-type effects in bilingual processing tasks: Support for a mixed-representational system. In R. Schreuder & B. Weltens (eds.), *The bilingual lexicon*, pp. 27–51. Amsterdam: John Benjamins.

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 3, 283–321.

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., Almeida, J., Janssen, N. & Caramazza, A. (in press). Lexical selection in bilingual speech production does not involve language suppression. *Journal of Experimental Psychology: Learning, Memory and Cognition*.

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., Nicol, J., Nakamura, K. & Greth, D. (2002). The role of language in memory for actions. *Journal of Psycholinguistic Research*, 31 (5), 447–457.

Glaser, W. R. & Glaser, M. O. (1989). Context effects on Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, 118, 13–42.

Gollan, T. H. & Acenas, L. A. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English bilinguals. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30, 246–269.

Gollan, T. H., Montoya, R. I., Fennema-Notestine, C., Morris, S. K. (2005). Bilingualism affects picture naming but not picture classification. *Memory & Cognition*, 33, 1220–1234.

Green, D. W. (1986). Control, activation and resource. *Brain and Language*, 27, 210–223.

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press

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 & Philadelphia: John Benjamins.

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1, 67–81.

Green, D. W. (2002). Representation and control: Exploring recovery patterns in bilingual aphasics. In F. Fabbro (ed.), *Advances in the neurolinguistics of bilingualism: Essays in honor of Michel Paradis*, pp. 239–259. Udine: Udine University Press.

Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology*, 93, 10–20.

Hermans, D. (2000). Word production in a foreign language. Ph.D. thesis, University of Nijmegan.

Hermans, D., Bongaerts, T., De Bot, K. & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition*, 1, 213–230.

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: Psycholinguistic approaches*, pp. 289–307. Oxford: 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, 1–75.

Li, P. & Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition*, 83, 265–294.

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. (1999). The selection of determiners in noun phrase production. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 25, 907–922.

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.

Morton, J. (1969). The interaction of information in word recognition. *Psychological Review*, 76, 165–178.

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: Lawrence Erlbaum Publishers.

Poulisse, N. & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics*, 15, 36–57.

Roelofs, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, 42, 107–142.

Schiller, N. O. & Caramazza, A. (2003). Grammatical feature selection in noun phrase production: Evidence from German and Dutch. *Journal of Memory and Language*, 48, 169–194.

Starreveld, P. A. & La Heij, W. (1996). Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 896–918.

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.

Received 5 November 2004
Revision received 15 September 2005
Accepted 25 September 2005

https://doi.org/10.1017/S1366728906002501 Published online by Cambridge University Press