Memory & Cognition
2002, 30 (4), 614-628

Stroop effects in bilingual translation

NATASHA A. MILLER and JUDITH F. KROLL
Pennsylvania State University, University Park, Pennsylvania

In two experiments, bilinguals proficient in English and Spanish translated words from one language to the other. In each experiment, following the target word to be translated, distractor words were presented after a short (200-msec) or long (500-msec) stimulus onset asynchrony. In Experiment 1, the distractor words appeared in the language of production and were related to the meaning or form of the spoken translation. The results replicated past studies in demonstrating that semantically related distractor words produced Stroop-type interference, whereas form-related distractor words produced facilitation. In Experiment 2, the distractors appeared in the language of input and were related to the meaning or form of the target word itself. In contrast to the results of Experiment 1, there were only marginal effects of the distractors on translation performance. These results suggest that language cues related to the nature of the input in translation may serve to reduce competition among lexical competitors during lexicalization. The contrast between these results and those in bilingual picture–word interference studies provides important constraints for models of language production and for claims about the locus of language selection.

To speak an idea or to name an object, an individual must first retrieve the concept associated with the idea or object and then specify the necessary phonology to articulate the word. In some instances, more than a single alternative name will be available, and the speaker will have to choose among the possible candidates. For example, Peterson and Savoy (1998) showed that when speakers named a picture of a couch, the alternative *sofa* was also activated to the point of having its phonology specified. What is remarkable about this finding is that it suggests that lexical alternatives may compete far into the process of lexicalizing concepts into words (see also Jescheniak & Schriefers, 1998). Although this scenario may be atypical when individuals speak within a single language, because few objects have close synonym names, for a proficient bilingual for whom almost every concept has a name in both languages, the translation equivalent of a word may routinely function as a close competitor.[1] Unless the intention to speak words in one language serves to selectively activate lexical candidates in the target language alone, the presence of a second language will have the consequence of increasing the degree of lexical competition during production.

Recent research on bilingual language production suggests that the intention to speak in one language does not eliminate activity of lexical alternatives in the other language, especially when the second language (L2) is spoken. That is, it does not appear that the bilingual can function as a monolingual. For example, Hermans, Bongaerts, de Bot, and Schreuder (1998) used the picture–word interference version of the Stroop task (Stroop, 1935) to investigate the activity of translation equivalents during picture naming in L2. In this task, a picture is presented for naming, and a distractor word precedes, follows, or is presented simultaneously with the picture. The participant is instructed to ignore the distractor while naming the picture. By varying the similarity of the distractor to the target's name and the timing of its presentation, one can infer the nature of the processes active prior to production. Distractors that are semantically related to the picture's name typically produce interference in picture naming, whereas distractors that are orthographically and/or phonologically similar to the picture's name typically produce facilitation (e.g., Lupker, 1979; Starreveld & La Heij, 1995).

Hermans et al. (1998) asked whether this pattern of results would be obtained when Dutch–English bilinguals named pictures in English with distractors that were spoken in either Dutch, the first language (L1), or in English, the L2. They found that like the earlier monolingual studies, there was reliable semantic interference and form facilitation. Moreover, the semantic interference was greatest at short stimulus onset asynchronies (SOAs), but the phonological facilitation was greatest at long SOAs. The presence of these effects when the distractor appeared in L1 but picture naming was performed in L2 suggests that

Portions of this research were presented at the 43rd Annual Meeting of the International Linguistic Association and the 70th Annual Meeting of the Eastern Psychological Association. This research was supported in part by NSF Grants BCS-9905850 and BCS-0111734 and by a sabbatical grant from the Nijmegen Institute for Cognition and Information, The Netherlands, to J.F.K. We thank Ton Dijkstra and an anonymous reviewer for comments on an earlier version of the manuscript and Michele Diaz for research assistance. Natasha Miller is now at the University of Maryland School of Medicine. Correspondence concerning this article should be addressed to J. F. Kroll, Department of Psychology, 641 Moore Building, Pennsylvania State University, University Park, PA 16802 (e-mail: jfk7@psu.edu).

—Accepted by previous editorial team

information from both languages was available for at least some time into the lexicalization process. In the critical condition of their experiment, Hermans et al. presented a distractor word in L1 or L2 that sounded like the name of the picture in L1. For example, if Dutch–English bilinguals named a picture of a mountain in English (*mountain*), the distractor word *berm* was presented, which sounds like *berg*, the Dutch name for mountain. The question was whether these "phono-Dutch" distractors would behave like semantic distractors or like phonological distractors. The results showed that they produced interference like the semantic distractors and also followed a similar time course to the semantic distractors, with the largest interference occurring at short SOAs. On the basis of these results, Hermans et al. argued that lexical alternatives in both of a bilingual's languages are active during production, but only to the level of selecting among abstract lexical forms, not to the level of the phonology. A similar series of experiments by Costa, Miozzo, and Caramazza (1999) provided converging evidence for the claim that lexical alternatives are available in both languages prior to naming a picture but that, unlike the monolingual case, they are not phonologically specified.

In the present study, we asked whether the same pattern of bilingual naming results would be observed when the production task is changed from picture naming to translation. The comparison between picture naming and translation has been used extensively in past studies of bilingual representation to evaluate the nature of the interlanguage connections for individuals at early stages of L2 learning and once they become proficient bilinguals (e.g., Chen & Leung, 1989; Kroll & Curley, 1988; Potter, So, von Eckardt, & Feldman, 1984). Although the two tasks share many of the same component processes, they also differ in ways that may potentially affect language selection. For example, in picture naming, the process of identifying the pictured object is thought to require the computation of a structural description that is likely to share depicted features with other objects drawn from the same semantic category (e.g., Lloyd-Jones & Humphreys, 1997). In translation, the presented word also has to be identified, but the processes that contribute to word recognition are likely to involve the activation of word neighbors that share lexical, rather than visual/semantic, features with the target. For words, lexical access precedes conceptual processing. For pictures, the process is reversed. However, in both tasks, once meaning is specified, the subsequent process of lexicalization is thought to be similar (but see Kroll & Stewart, 1994).

In one past study, La Heij, de Bruyn, Elens, Hartsuiker, and Helaha (1990) examined Stroop-type effects in a translation task in which Dutch–English bilinguals performed translation from L2 (English) to L1 (Dutch). Distractor words were presented visually either 400 msec before or 140 msec after an English word to be translated. La Heij et al. found no effects at the SOA of −400 msec, but at 140 msec there was interference when distractor words were semantically related and facilitation when distractors were orthographically related to the translation. The pattern at the positive SOA was thus similar to the one reported for picture–word interference.

In Experiment 1, we first replicated the main conditions of La Heij et al. (1990) with bilinguals who speak English and Spanish. We included two different positive SOAs (200 and 500 msec) and semantically related and form-related distractors, as well as unrelated controls, and bilinguals performed both directions of translation (i.e., forward translation from L1 to L2 and backward translation from L2 to L1). As in La Heij et al., the distractors were presented in the language of output (i.e., in L2 when participants translated from L1 to L2 and in L1 when they translated from L2 to L1). We predicted that there would be similar patterns of semantic interference and lexical form facilitation but that the semantic effects might be larger at the short SOA and the form effects larger at the long SOA (e.g., Levelt, Roelofs, & Meyer, 1999). Because production in L2 tends to be slower and more error prone, owing to the consequences of lexical competition, we also expected that the magnitude of the effects would be larger in the forward translation task.

In Experiment 2, we introduced a new manipulation that is critical for determining the locus of selection effects in bilingual production. The conditions in this experiment were identical to those in Experiment 1, except that distractor words were now presented in the language of the input, rather than in the language of the output. For example, if a bilingual was asked to translate *cat* into *gato*, the word *dog* would be a semantically related distractor, and the word *can* would be a form-related distractor. If semantically related lexical alternatives are active in both languages prior to selecting the word to be spoken, the language of the distractor should have little impact on the presence of interference effects; they should be similar regardless of the language in which the semantic information is presented. The results of cross-language picture–word interference experiments suggest that this is the case. Both Hermans et al. (1998) and Costa et al. (1999) found similar magnitudes of semantic interference in picture naming, regardless of whether semantically related distractors were presented in the target or the nontarget language for production. However, if the information in the input to translation specifies the language of production (i.e., when a word is presented in L1 for translation, the bilingual knows immediately that he or she is to produce in L2, and not in L1), only semantic alternatives in the output may function as competitors. Our hypothesis is that one important way in which translation differs from picture naming is in the specification of a language cue at the moment the input word is presented. According to this view, only semantically related distractors in the output language will produce interference. For form distractors, words that resemble the stimulus input should increase the level of competition among lexical neighbors and, therefore, produce interference, rather than the facilitation observed when distractors appear in the language of output.

To anticipate the findings, the results of Experiment 1, with distractors in the language of output, closely resembled those reported by La Heij et al. (1990) for translation and by Hermans et al. (1998) and Costa et al. (1999) for picture naming. In contrast, in Experiment 2, with distractors in the language of the input, we found no effects of either distractor type. In the General Discussion section, we will consider the implications of the pattern of results for models of bilingual production and, in particular, for the issue of whether there is a fixed locus of language selection.

EXPERIMENT 1
Distractors in the Language of Production

In Experiment 1, a word was presented visually in either English or Spanish, and participants were asked to translate it as quickly as possible into the other language. The word was followed, after a short (200-msec) or long (500-msec) SOA, by a distractor word to be ignored. The distractor word either was semantically or form related to the word to be spoken or was an unrelated control. The main question in Experiment 1 was whether we could replicate the pattern of inhibition for semantic distractors and facilitation for lexical form distractors reported by La Heij et al. (1990).

Method

Materials

Ninety-six English words and their Spanish translations were selected as targets. Items were selected to include words of high and low frequency and different grammatical classes. All the target words were noncognate translations.

Half of the words (48) were presented as English targets, and the other half (48) were presented as Spanish targets. Forty additional words, 20 English and 20 Spanish, were included as practice targets. The average English word frequency for target words in English to Spanish and Spanish to English translation was 122 and 126 times per million, respectively (Francis & Kučera, 1982). The average word length for these conditions was 5.3 and 4.8 letters, respectively. The complete materials are available from the authors.

Distractor words appeared in the language of the correct translation and were semantically related to the correct response, unrelated controls for the semantic condition, form related to the correct response, or unrelated controls for the form condition. For example, for the translation *cuchara* to *spoon*, the related form distractor was *spool* and the related semantic distractor was *fork*. The two unrelated controls were formally similar. Both were included to provide a closer match to the critical distractors on lexical properties than would have been possible with a single unrelated condition. In addition, a row of *x*s served as a neutral baseline condition. Distractors were chosen so that the mean word length and word frequency were similar across related and unrelated conditions. Form-related distractors included words that shared phonology and/or orthography. Semantically related distractors included words that were exemplars of the same category or words that were semantically associated. The mean word frequency for form-related and form-unrelated words was 106 and 116 times per million, respectively. The mean word length for these conditions was 5.0 and 5.1 letters, respectively. The mean frequency for semantically related and unrelated words was 114 and 112 times per million, respectively. The mean word length for these conditions was 5.2 and 5.3 letters, respectively. In the neutral condition, the number of *x*s in the distractors matched the average number of letters in the critical conditions. A series of t tests failed to reveal any differences between the related and the unrelated words (p values $>.05$).

Design and Procedure

The independent variables were direction of translation (L1 to L2 or L2 to L1), SOA (200 or 500 msec), distractor type (semantic, form, and neutral), and the relatedness of the semantic and form distractors (related to unrelated). SOA was manipulated between groups. All other factors were manipulated within groups. Different versions of the materials were created so that, for any given participant, a target would appear only once in one of the distractor conditions. The participants were randomly assigned to one SOA condition and to one of the versions. The order of materials within a list was randomized, and the order of the translation tasks was counterbalanced across participants.

The experiment was performed on an IBM computer. Instructions and a practice block preceded the experimental trials. To begin the task, the participants were instructed to press any key to remove the instructions from the screen. The instructions were then replaced with a cross (+) in the middle of the screen. Upon pressing a key, the cross was replaced by the target word, followed by the distractor in the same location. In the short-SOA condition, the target word was presented on the computer screen for 200 msec, the interstimulus interval (ISI) was 0 msec, and the distractor then appeared on the screen for 100 msec. In the long-SOA condition, the ISI was extended to 300 msec. The participants were required to speak aloud the correct translation. If they did not know the correct translation, they were instructed to respond "no" or "I don't know." Following a response, the cross reappeared on the screen to indicate the start of the next trial. Spoken responses were tape recorded so that a proficient bilingual could later code accuracy.

In the final phase of the experiment, the participants completed a language history questionnaire to obtain information about their language experience and self-assessed proficiency in writing, reading, conversation, and comprehension in each of their languages. The duration of the experiment was approximately 40 min.

Participants

Fifteen English/Spanish and 20 Spanish/English bilinguals at Pennsylvania State University participated in the experiment. Ratings from the language history questionnaire revealed that the participants generally rated themselves as highly proficient in both languages. The profile of their language skills is shown in Table 1. To determine language dominance, a mean rating was computed for each participant on the basis of the four self-assessed skills. The language with the highest mean rating was identified as the dominant language. If both languages were rated equally, English was identified as the dominant language on the assumption that these individuals were immersed in an English-speaking environment. Using this procedure, seven English-dominant participants were identified from the sample of native Spanish speakers. The data of 3 participants were excluded from the analysis because they failed to meet an overall accuracy criterion of 80%. Of the final sample of 32 participants, 19 were identified as English dominant, and 13 were identified as Spanish dominant.

Results and Discussion

Data Analysis

The data were analyzed to examine correct mean reaction times (RTs) and percentages of accuracy in the translation task. Data on trials in which RTs were below 300 msec and above 3,000 msec were considered outliers

Table 1
Characteristics of Participants in Experiments 1 and 2, Including Mean Self-Assessed Proficiency Ratings on Reading, Writing, and Conversation in L2

	Experiment 1		Experiment 2	
Characteristic	Short SOA	Long SOA	Short SOA	Long SOA
Mean age (in years)	24.6	26.8	24.6	23.9
Number of years in U.S.	13.8	12.1	15.3	15.9
Rating of reading in L1	9.7	10.0	9.6	9.7
Rating of reading in L2	8.4	8.8	8.7	8.3
Rating of writing in L1	9.5	9.8	9.1	9.5
Rating of writing in L2	7.8	8.1	8.1	7.8
Rating of L1 conversational ability	9.9	9.8	9.7	9.9
Rating of L2 conversational ability	7.9	8.4	8.3	8.0
Rating of L1 comprehension	9.8	9.5	9.9	9.7
Rating of L2 comprehension	9.3	9.5	8.9	8.7
Mean L1 ratings	9.7	9.8	9.6	9.7
Mean L2 ratings	8.3	8.4	8.4	8.2

Note—Each scale was rated from 1 to 10, where 1 was *not very proficient* and 10 was *highly proficient.* SOA, stimulus onset asynchrony; L1, first language; L2, second language.

and were scored as errors. A trial was also scored as an outlier if the RT was greater than 2.5 standard deviations above the mean of correct responses for a given participant. Outliers accounted for 7% of the data. Because the participants in Experiment 1 varied in their language dominance, the data were coded in terms of the appropriate L1 and L2. For example, if an individual was coded as Spanish dominant, Spanish was considered the L1, and English was considered the L2. All subsequent analyses were conducted with the language variable coded as L1 and L2.

Because language dominance was determined following the assignment of participants to SOA groups, an additional set of analyses was performed to determine whether the SOA groups were otherwise similar in terms of language proficiency. A series of *t* tests was performed on each of the rating scales within the language history questionnaire to determine whether the two SOA groups differed. The analyses revealed no significant differences between the two SOA conditions on the majority of the rating scales (p values $> .05$). The single significant difference that emerged between the two SOA groups was in their ratings of their proficiency in reading in L1 [$t(30) = -2.61$, $p < .05$]. The short-SOA group rated their English reading skills as 9.5, and the long-SOA group rated theirs as 9.8, on a 10-point scale. It seems likely that this difference was significant because of the low variability of L1 ratings close to the ceiling of the scale.

Neutral Distractor Trials

RTs and accuracy to perform translation in the neutral condition, where *x*s matched in length to the target replaced actual word distractors, were analyzed to examine the effects of direction of translation and SOA. Analyses on the participants' neutral data included one between-subjects factor (SOA, 200 vs. 500 msec) and one within-subjects factor (direction of translation L1 to L2 vs. L2 to L1). The analysis by items was conducted with both SOA and direction of translation as within-items factors.

Reaction time. Mean correct translation RTs for the neutral distractor trials across both SOA conditions and both directions of translation are reported in Table 2. A translation asymmetry was observed, with longer RTs for L1 to L2 translation than for L2 to L1 translation. Furthermore, RTs were longer in the long-SOA than in the short-SOA condition. Analyses of variance (ANOVAs) revealed a significant main effect of direction of translation for participants [$F_1(1,30) = 6.21$, $MS_e = 16{,}172.5$, $p < .025$] and for items [$F_2(1,95) = 5.39$, $MS_e = 53{,}006.4$, $p < .025$]. A main effect of SOA was also observed by participants, with longer RTs in the long-SOA relative to the short-SOA condition [$F_1(1,30) = 4.67$, $MS_e = 41{,}300.1$, $p < .05$] and by items [$F_2(1,95) = 20.94$, $MS_e = 48{,}127.8$, $p < .01$]. The interaction between direction of translation and SOA was not significant in either analysis (p values $> .10$).

Accuracy. The percentages of accuracy for the neutral distractors trials are also reported in Table 2. These data

Table 2
Mean Translation Latencies (in Milliseconds) and Accuracy for Neutral Distractor Trials for Both Stimulus Onset Asynchrony (SOA) Conditions in Experiments 1 and 2

	Direction of Translation			
	L1 to L2		L2 to L1	
SOA Condition	*M*	%	*M*	%
	Experiment 1			
Short (200 msec)	1,219	83	1,181	84
Long (500 msec)	1,370	83	1,250	85
	Experiment 2			
Short (200 msec)	1,364	79	1,365	71
Long (500 msec)	1,380	75	1,387	74

Note—L1, first language; L2, second language.

suggest that accuracy was relatively high across all of the neutral conditions (the mean was approximately 83%). None of the main effects or interactions was significant for either the analysis by participants or the analysis by items (all p values $> .05$). Because accuracy was relatively high and similar across all conditions, these data suggest that the between-subjects manipulation of SOA resulted in bilingual groups that were equally proficient in their two languages.

The longer RTs for L1 to L2 than for L2 to L1 translation replicate the translation asymmetry reported in previous studies (e.g., Kroll & Stewart, 1994; Sánchez-Casas, Davis, & García-Albea, 1992; Sholl, Sankaranarayanan, & Kroll, 1995). According to the revised hierarchical model proposed by Kroll and Stewart, the asymmetry reflects a difference in the component processes engaged by the two translation tasks. Forward translation, from L1 to L2, is hypothesized to require conceptual access. Backward translation, from L2 to L1, is hypothesized to be accomplished on the basis of lexical associations between the L2 word and its translation equivalent in L1. Although the magnitude of the asymmetry was somewhat larger for the long-SOA than for the short-SOA group, the failure to observe a significant interaction suggests that the difference was not reliable.

The similar accuracy for the two SOA groups suggests that the RT difference that was observed is not likely to be due to a difference in proficiency between these two groups. The observed SOA difference may be due to a cuing effect induced by the fact that SOA was a blocked variable. The participants may have adopted a response criterion that was locked to the timing of the distractor presentation. Recent studies suggest that the temporal properties of trial events may influence RTs (e.g., Taylor & Lupker, 2001).

Critical Semantic and Form Distractor Trials

ANOVAs on RTs from critical distractor trials were performed with one between-subjects factor (SOA, 200 vs. 500 msec) and three within-subjects factors—direction of translation (L1 to L2 vs. L2 to L1), type of distractor (form vs. semantic), and relatedness of the distractor (related vs. unrelated). In the analyses by item, SOA, direction of translation, type of distractor, and relatedness of distractor were all within-items factors.

Reaction time. The mean RTs and accuracy for semantically related and form-related distractors and their controls are given in Table 3 for both SOA groups and for both directions of translation. Overall, the results show that, like previous studies (e.g., Glaser & Düngelhoff, 1984; La Heij et al., 1990), semantic distractors tended to produce interference, whereas form distractors produced facilitation.

To determine the extent to which the direction of translation and SOA condition modulated this general pattern, ANOVAs were again performed by participants and by items. The analysis by participants revealed a significant interaction between the type and the relatedness of the distractor [$F_1(1,30) = 17.29$, $MS_e = 12,496.1$, $p < .01$], however, the same interaction was only marginally significant in the analysis by items [$F_2(1,95) = 2.97$, $MS_e = 112,008.7$, $p < .10$]. This interaction provides support for the claim that semantic distractors produced interference ($M = 1,400$ msec for semantically related distractors and $M = 1,340$ msec for unrelated controls), whereas form distractors produced facilitation ($M = 1,294$ msec for form-related distractors and $M = 1,349$ msec for unrelated controls). Simple effects tests showed that the relatedness effect was significant for both the form and the semantic conditions [$F(1,30) = 5.38$, $p < .05$, and $F(1,30) = 8.47$, $p < .01$, respectively].

Table 3
Mean Response Latencies (in Milliseconds) and Accuracy to Translate Words in Both Directions of Translation as a Function of the Stimulus of Asynchrony (SOA), Type of Distractor (Form or Semantic), and Relatedness of the Distractor to the Word to be Spoken (Related or Unrelated) in Experiment 1

Language Condition	Relatedness: Related *M*	Related %	Unrelated *M*	Unrelated %	Difference (Related − Unrelated)
			L1–L2		
Short SOA					
Form	1,319	79	1,371	80	−52
Semantic	1,468	78	1,353	77	115
Long SOA					
Form	1,433	82	1,423	70	10
Semantic	1,467	70	1,412	74	55
			L2–L1		
Short SOA					
Form	1,163	84	1,268	81	−105
Semantic	1,332	77	1,278	84	54
Long SOA					
Form	1,260	84	1,336	88	−76
Semantic	1,333	85	1,316	83	17

Although the four-way interaction between SOA, direction of translation, type of distractor, and relatedness was not significant by participants or by items, it is informative to examine the data across these conditions. To get a clearer picture of how the form and semantic distractors affected translation in these conditions, we computed a difference score between the related and the unrelated distractors and plotted the magnitude of interference and facilitation for each SOA and for both directions of translation and both distractor types. These graphs are shown in Figure 1. Figure 1A gives the magnitude of semantic interference observed for both SOAs and directions of translation. Figure 1B gives the magnitude of form facilitation for the same conditions.

The interaction between direction of translation and relatedness of the distractor was significant in the analysis by participants [$F_1(1,30) = 4.41$, $MS_e = 12{,}633.1$, $p < .05$], but not in the analysis by items [$F_2(1,95) = 2.36$, $MS_e = 71{,}114.2$, $p > .05$]. A marginally significant interaction between SOA and type of distractor was also obtained, but only in the analysis by participants [$F_1(1,30) = 4.05$, $MS_e = 13{,}698.7$, $p < .06$], and not in the analysis by items [$F_2 > .05$]. Analyses on both participant and item RTs revealed a significant main effect of direction of translation by participants [$F_1(1,30) = 10.87$, $MS_e = 84{,}784.6$, $p < .05$] and by items [$F_2(1,95) = 19.78$, $MS_e = 75{,}547.5$, $p < .01$]. Translation from L1 to L2 resulted in longer RTs than did L2 to L1 translation. Like the results reported for the neutral trials, this finding provides support for the translation asymmetry predicted by the revised hierarchical model.

There was also a significant main effect of type of distractor in the analysis by participants, so that form-related distractors produced faster RTs than did semantically related distractors overall [$F_1(1,30) = 10.93$, $MS_e = 13{,}698.7$, $p < .05$], and a marginally significant effect in the analysis by items [$F_2(1,95) = 3.73$, $MS_e = 99{,}658.2$, $p < .06$].

Although RTs in the long-SOA condition were significantly longer than those in the short-SOA condition for the neutral trials, in the context of the critical distractor trials, the SOA difference was not significant for participants [$F_1(1,30) = 0.888$, $MS_e = 206{,}246.4$, $p > 05$] but was significant for items [$F_2(1,95) = 11.21$, $MS_e = 136{,}870.7$, $p < .01$]. It is possible that the failure to obtain significant interactions with SOA in this experiment was due to the choice of SOA values or to the use of a blocked SOA design.

Overall, the results closely replicate those reported by La Heij et al. (1990). In the present study, there is a suggestion that the form effects are larger at the shorter SOA than at the longer SOA, contrary to the results in cross-language picture–word interference (Hermans et al., 1998). However, the interaction between SOA, type of distractor, and relatedness was only marginally significant [$F_1(1,30) = 2.83$, $MS_e = 12{,}496.1$, $p = .11$]. The form facilitation that was observed may be a result of the distractor's facilitating the selection of the appropriate word. This effect may be most likely to occur in the short-SOA condition because the semantic alternatives were still active. This idea has been suggested in previous monolingual studies in which phonological facilitation has been observed at short SOAs. The claim is that phonological effects can arise in two ways, either by reflecting a later stage of phonological encoding or by serving to reduce competition among lexical alternatives at an earlier stage (see Starreveld, 2000).

Accuracy. An ANOVA revealed a four-way interaction between SOA, direction of translation, type of distractor, and relatedness of distractor in the analysis by participants [$F_1(1,30) = 4.77$, $MS_e = 0.025$, $p < .05$] and a marginally significant effect in the analysis by items [$F_2(1,95) = 3.02$, $MS_e = 0.063$, $p < .09$]. A main effect of direction of translation was obtained in the analysis by participants [$F_1(1,30) = 11.98$, $MS_e = 0.028$, $p < .05$]; however, the interaction was not significant in the analysis by items [$F_2(1,95) = 0.495$, $MS_e = 0.147$, $p > .05$]. Replicating previous translation studies (e.g., Kroll & Stewart, 1994), accuracy was higher from L2 to L1 than from L1 to L2. Further analysis revealed a marginally significant effect of distractor type for participants [$F_1(1,30) = 2.98$, $MS_e = 0.014$, $p < .10$], but not for items [$F_2(1,95) = 2.65$, $MS_e = 0.082$, $p > .05$]. Higher accuracy was observed for form-related distractors than for semantically related distractors. An analysis of the items revealed no further effects.

Relation to Previous Findings

The results replicated the main findings of semantic interference and form facilitation reported by La Heij et al. (1990) for L2 to L1 translation and extended them to L1 to L2 translation. Although the interaction between direction of translation and the magnitude of the semantic interference effect was not significant, the data shown in Figure 1 are consistent with the prediction of the revised hierarchical model that, if anything, semantic effects should be larger in the L1 to L2 direction, hypothesized to be conceptually mediated. Likewise, although the effect of direction of translation did not interact significantly with the magnitude of form facilitation, there is a suggestion in the data that, if anything, these effects are larger in the L2 to L1 task, the direction of translation hypothesized to be lexically mediated. The finding that there was semantic interference in the L2 to L1 direction of translation particularly raises problems for the revised hierarchical model. Because the model assumes that L2 to L1 translation can be performed lexically without semantic mediation, the presence of semantic interference in this direction is potentially damaging evidence (see Kroll & de Groot, 1997).

The results of Experiment 1 are also consistent with the findings of the picture–word interference studies reported by Hermans et al. (1998) and Costa et al. (1999). Semantic interference and form facilitation were observed in both directions of translation when the distractor was presented in the language of the spoken response. These results provide converging support for the hypothesis that

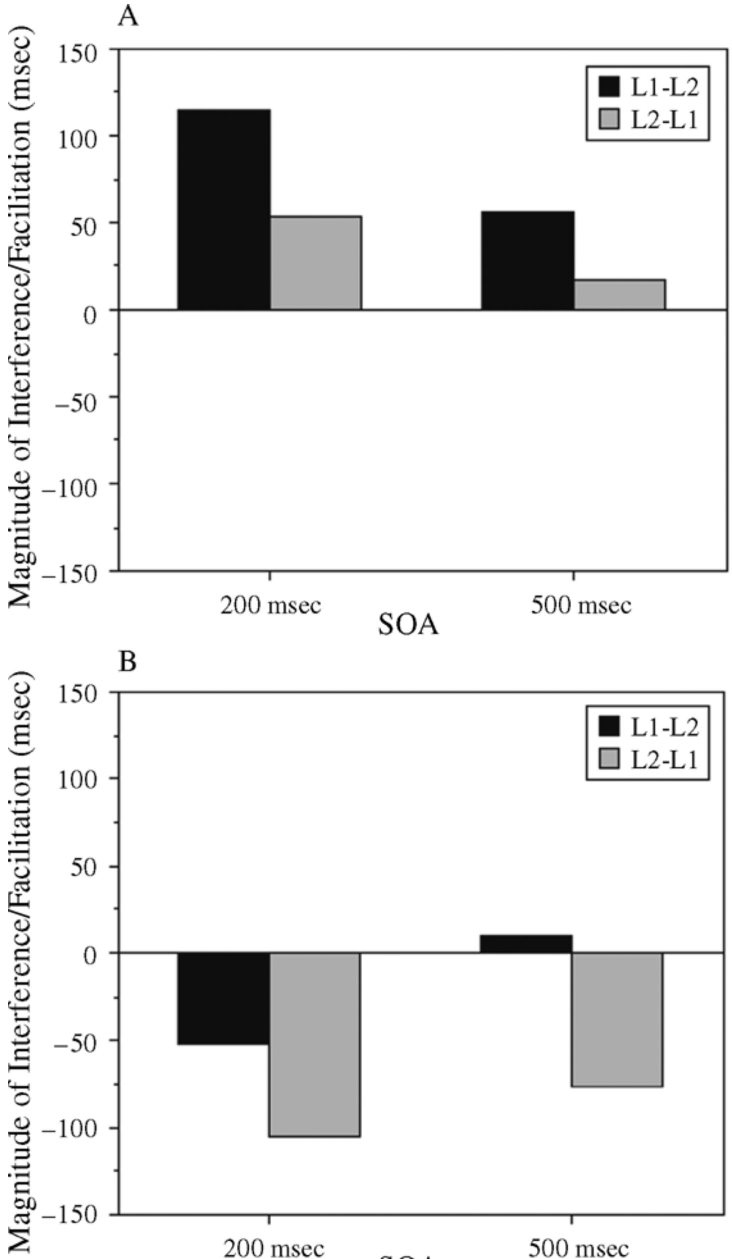

Figure 1. The magnitude of facilitation (negative) and interference (positive) in milliseconds for (A) semantically related distractors relative to controls and (B) form-related distractors relative to controls as a function of stimulus onset asynchrony (SOA, 200 vs. 500 msec) and direction of translation (L1 to L2 vs. L2 to L1) in Experiment 1.

picture naming and translation engage similar component processes.

EXPERIMENT 2
Distractors in the Language of Input

Experiment 2 was identical to Experiment 1, with a single exception: The language of the distractor word now matched the language of the target input, rather than the language of the translation to be spoken. If semantically related alternatives in both languages compete prior to selection of the word to be spoken, the language of the distractor should have little impact on the presence of interference effects; any factor that increases the degree of competition should produce interference. However, if the language of input in translation provides a cue that allows language selection to occur relatively early in production, only alternatives in the language of production should compete, and distractors in the language of input should have little effect.

In Experiment 2, the form-related distractors resembled the input target word. Because the form distractors included words that were related both orthographically and phonologically to the target words, we expected that the effect of the form distractors would be to increase competition at the level of recognizing the target word. In contrast to the results of Experiment 1, form distractors in Experiment 2 were predicted to produce interference, rather than facilitation.

Method

Materials

The same English and Spanish words and their translations as those from Experiment 1 were used in Experiment 2, with one exception. Words that were used in Experiment 1 for Spanish to English translation were now used for English to Spanish translation (e.g., in Experiment 1, *lluvia* was translated into *rain*, whereas in Experiment 2, *rain* was translated into *lluvia*). This change was necessary because the distractors were now related to the input target, rather than to the spoken translation.

Design and Procedure

The design and procedure were identical to those described for Experiment 1, with the exception that the distractor words were now related to the target input.

Participants

The participants were 17 English/Spanish and 19 Spanish/English bilinguals from Pennsylvania State University, who had not participated in Experiment 1. Because Experiment 2 was conducted to test predictions based on the results of Experiment 1, it was important that the same criteria be used to determine language proficiency and dominance. Table 1 gives a profile of the participants' language skills. Ratings from the language history questionnaire revealed that the participants generally rated themselves as highly proficient in both languages. As in Experiment 1, these ratings were used to determine language dominance. By using the procedure described earlier, 2 of the native Spanish speakers were identified as English dominant. Data from 5 of the 36 participants were excluded because they failed to meet an accuracy criterion of 80% across all conditions. Of the final sample of 31 participants, 19 participants were identified as English dominant, and 12 participants were identified as Spanish dominant.

A series of analyses was performed to compare the proficiency of the participants in the two experiments and in each of the two SOA conditions within each experiment. The mean overall ratings for L1 and L2 are shown in Table 1 for both experiments. ANOVAs on the mean ratings for each language revealed no main effects of either between-group variable (experiment or SOA), and no signficant interaction between them (all p values $> .25$). Similar analyses on mean age and number of years in the U.S. revealed no reliable main effects or interactions (all p values $> .25$). These comparisons suggest that the participants across the between-group conditions of the two experiments were well matched.

Results and Discussion

Data Analysis

The data were analyzed to examine the correct mean RTs and percentages of accuracy. Data on trials in which RTs were below 300 msec and above 3,000 msec were considered outliers and were scored as errors. A trial was also scored as an outlier if the RT was greater than 2.5 standard deviations above the mean for a given participant. Outliers accounted for 6% of the data. As in Experiment 1, the data were coded in terms of the dominant L1 and L2.

Neutral Distractor Trials

RTs and accuracy in the neutral condition were analyzed to examine the effects of direction of translation and SOA. Analyses on the participants' neutral data included one between-subjects factor (SOA, 200 vs. 500 msec), and one within-subjects factor (direction of translation, L1 to L2 vs. L2 to L1). The analysis by items was conducted with both SOA and direction of translation as within-subjects factors.

Reaction time. Mean correct translation RTs for the neutral distractor trials across both SOA conditions and both directions of translation are reported in Table 2. Unlike the results for Experiment 1, no translation asymmetry was observed. ANOVAs conducted by participants and by items failed to reveal a significant effect of translation direction or an interaction between translation direction and SOA (all p values $> .05$). A significant effect of SOA was observed for items [$F_2(1,95) = 6.58$, $MS_e = 69{,}168.5, p < .05$]. Like the results of Experiment 1, RTs were longer in the long-SOA than in the short-SOA condition.

Because the participants in the two experiments were closely matched, the elimination of the translation asymmetry in Experiment 2 seems most likely to be due to the presence of same-language distractors. Recent studies on bilingual word recognition (e.g., Van Heuven, Dijkstra, & Grainger, 1998) have demonstrated that lexical form relatives are activated in both of a bilingual's languages even when a word has to be recognized in one language only. In Experiment 2, the presence of distractors in the same language as the target input may have served to increase the number of activated lexical competitors. The effects of this competition should be greatest when the distractors were lexical form relatives of the target. Although the neutral trials were identical in both Experiments 1 and 2 (i.e., rows of *x*s that were neither in the language of the input target nor in the language of the spoken output), the difference in the context in which they were embedded in the two experiments may be responsible for the observed effects.

Why would increased lexical competition from same-language distractors eliminate the translation asymmetry? We hypothesize that the effects of lexical competition will be greatest when L2 target words are followed by L2 distractors because the less dominant L2 is likely to be processed more slowly than L1 and, therefore, the process of binding of the L2 target word to its meaning is less likely to be complete at the point at which the distractor appears. Thus L2 to L1 translation will be slowed relative to L1 to L2 translation, and the translation asymmetry will disappear. In the case of L2 to L1 translation, in particular, competition between the distractor and the target may also have the consequence of creating potential confusion about which word is to be translated. The higher error rate in Experiment 2, coupled with longer RTs, supports this interpretation.[2]

Accuracy. No effects were significant in the analyses performed on the accuracy data. However, as compared to the neutral condition of Experiment 1, the accuracy in Experiment 2 was considerably lower. As was suggested earlier, the reduced accuracy in Experiment 2 seems more likely to be due to the same language distractors, rather than to differences among the bilingual groups in these conditions.

Critical Semantic and Form Distractor Trials

As in Experiment 1, ANOVAs on RTs from critical distractor trials were performed with one between-subjects factor (SOA, 200 vs. 500 msec) and three within-subjects factors—direction of translation (L1 to L2 vs. L2 to L1), type of distractor (form vs. semantic), and relatedness of the distractor (related vs. unrelated). In the analyses by item, SOA, direction of translation, type of distractor, and relatedness of distractor were all within-items factors.

Reaction time. The mean RTs for semantically related and form-related distractors and their controls are given in Table 4 for both SOA groups and for both directions of translation. Table 4 shows that the pattern of results in Experiment 2 was quite different from the one obtained in Experiment 1 and by La Heij et al. (1990) when distractors appeared in the language of production. The semantic interference that was observed in both directions of translation in Experiment 1 is now reduced, and there is even a suggestion of semantic facilitation in the L1 to L2 direction at the long SOA. The pattern for the form distractors is also quite different from the one observed in Experiment 1. Whereas facilitation was obtained in almost all the conditions of Experiment 1 (with the exception of L1 to L2 translation at the long SOA), in Experiment 2 there was facilitation only for L1 to L2 translation. For L2 to L1 translation, there was an interference effect at the short SOA, and no effect at all for the form distractors at the long SOA.

ANOVAs were performed to evaluate the reliability of these patterns. The analysis revealed a marginally significant interaction between direction of translation and relatedness of the distractor by participants [$F_1(1,29) = 3.76$, $MS_e = 34{,}874.2$, $p < .10$] and by items, [$F_2(1,95) = 3.50$, $MS_e = 81761.1$, $p < .10$]. Overall, longer RTs were obtained in L2 to L1 translation when the distractors were related to the target. Although no other significant effects were observed in the analysis by participants, the analysis by items revealed a marginally significant interaction between SOA and direction of translation [$F_2(1,95) = 2.86$, $MS_e = 97{,}147.3$, $p < .10$].

Accuracy. An analysis revealed a significant interaction between SOA and the relatedness of the distractor by participants [$F_1(1,29) = 6.80$, $MS_e = 0.014$, $p < .05$] and a marginal interaction by items [$F_2(1,95) = 4.03$, $MS_e = 0.096$, $p < .10$]. A marginal interaction was also observed between SOA and type of distractors for both participants [$F_1(1,29) = 3.28$, $MS_e = 0.118$, $p < .10$] and items [$F_2(1,95) = 3.48$, $MS_e = 0.106$, $p < .10$]. No other significant effects were obtained on the basis of the participant data; however, the analysis by items revealed a marginal interaction between SOA and direction of translation [$F_2(1,95) = 2.86$, $MS_e = 0.078$, $p < .10$].

These results demonstrate that the semantic interference and form facilitation observed in Experiment 1, in which the distractors appeared in the language of production, were eliminated when the distractors appeared in the language of the target. At a general level, the differences between the two experiments suggest that the observed semantic interference and form facilitation in Experiment 1

Table 4
Mean Response Latencies (in Milliseconds) and Accuracy for Translating Words in Both Directions of Translation as a Function of Stimulus Onset Asynchrony (SOA), Type of Distractor (Form or Semantic), and Relatedness of the Distractor to the Input Word (Related or Unrelated) in Experiment 2

Language Condition	Relatedness: Related *M*	Related %	Unrelated *M*	Unrelated %	Difference (Related − Unrelated)
			L1–L2		
Short SOA					
Form	1,403	74	1,441	78	−38
Semantic	1,428	74	1,426	74	2
Long SOA					
Form	1,430	71	1,514	73	−84
Semantic	1,419	76	1,485	74	−66
			L2–L1		
Short SOA					
Form	1,510	71	1,427	80	83
Semantic	1,433	68	1,402	70	31
Long SOA					
Form	1,410	75	1,406	70	4
Semantic	1,447	79	1,383	70	64

Note—L1, first language; L2, second language.

and in previous reports in the literature (e.g., Starreveld & La Heij, 1995) can be attributed to the activation of lexical alternatives within one of the bilingual's languages. Only when the lexical candidates are potential responses do these effects appear.

Comparing Experiments 1 and 2. Although the analyses of the data in Experiment 2 failed to reveal significant main effects or interactions, we conducted a set of analyses to compare the two experiments more directly. As was stated previously, because the bilingual participants were carefully matched, it is unlikely that the between-group differences in the two experiments can be attributed to differential language proficiency. Instead, they appear to be attributable to a change in processing induced by the language of the distractors. The neutral trials provide a particularly good way to test this hypothesis, because the language of the distractors was irrelevant on these trials. Any residual effects on the neutral trials presumably reflect processing strategies that were created by the context in which these trials were embedded.

Analyses on the participant data were performed to compare the two experiments directly. The analysis of neutral RTs revealed a marginally significant interaction between experiment and direction of translation [$F_1(1,59) = 2.86$, $MS_e = 18990.9$, $p < .10$] and a main effect of experiment [$F_1(1,59) = 6.53$, $MS_e = 68083.9$, $p < .05$], with longer RTs in Experiment 2 than in Experiment 1 (1,255 vs. 1,474 msec, respectively). The same analysis on the accuracy data also revealed a main effect of experiment [$F_1(1,59) = 12.47$, $MS_e = 0.02$, $p < .01$], with higher accuracy in Experiment 1 than in Experiment 2 (83.7% vs. 74.7%, respectively). There were no other significant effects in either analysis (p values $> .05$).

A similar set of analyses was performed for the participant data on critical distractor trials. In the RT analysis, there was a significant interaction between experiment, type of distractor, and distractor relatedness [$F_1(1,59) = 4.08$, $MS_e = 19{,}508.3$, $p < .05$]. In Experiment 1, form-related distractors provided facilitation, whereas semantically related distractors produced interference; however, in Experiment 2, these effects were absent. These data are shown in Figure 2. Figure 2A gives the magnitude of semantic and form effects for both SOAs when the distractor appeared in the language of output in Experiment 1. Figure 2B shows the same conditions when the distractor appeared in the language of input in Experiment 1. Newman–Keuls tests showed that, in Experiment 1, the semantic interference effect was significant [$q(3,59) = 3.43$, $p < .05$] but that the form facilitation effect just missed significance [$q(3,59) = 3.15$, $p > .05$]. A two-tailed paired t test on the form effect was reliable [$t(31) = -2.323$, $p < .05$]. The same effects in Experiment 2 were not significant either by Newman–Keuls test or by one-tailed t tests (all p values $> .05$). The overall pattern in this interaction reflects the main difference between the two experiments. There were effects of form facilitation and semantic interference only when the distractors appeared in the language of the output.[3]

GENERAL DISCUSSION

In two experiments, we examined the nature of Stroop-type interference in bilingual translation. The results of Experiment 1, in which word distractors appeared in the language of production, replicated the previous findings of La Heij et al. (1990), in that semantically related distractors produced interference and form-related distractors produced facilitation, relative to unrelated controls. However, the results of Experiment 2, in which distractor words appeared in the language of the to-be-translated target, failed to reveal consistent effects of the distractor type on performance.

The results of Experiment 2 suggest that the cause of the semantic interference observed in Experiment 1 is the increased competition between lexical entries within the language of production. When the distractor is no longer a member of the production set (i.e., no longer a viable candidate for output), semantic interference is eliminated. Although the results reported here are consistent with the notion of a response set, they do not constrain the manner in which a response set might operate (see Levelt et al., 1999). In particular, because we failed to observe a clear pattern of SOA effects in these experiments, it is difficult to identify the locus of this mechanism.

Production Errors

Additional support for the hypothesis that semantic interference in the translation-Stroop task arises from lexical competition within one language alone comes from an analysis of the form of errors that were produced in these experiments. In most production studies, materials are repeated across trials, and when picture targets are used, they are typically preexposed to participants to reduce variability in assigning picture naming and to maximize the number of correct trials. In the present experiments, neither word targets nor distractors were repeated for a given participant. Rather, materials were counterbalanced across participants and conditions. The advantage of this procedure is that it provides insight into the processes that are engaged when the bilingual speaker has not already been primed to select a particular response alternative. However, it has the consequence that participants make errors, either because they do not know the translation for a given word or because they produce the incorrect translation. Thus, like research on naturally occurring speech errors (e.g., Poulisse, 1997), it is possible to examine the form of errors in the translation task as a source of converging evidence.

We examined two types of errors that the participants made during translation. Although these errors did not occur with high frequency, they occurred often enough for us to ask whether they were related to the nature of the distractor condition. One type of error was the mistake of naming the distractor itself. For example, if the bilingual participant was translating the word *dinero* and the distractor was the word *dollar*, we asked whether he or she ever made the mistake of saying *dollar* rather than *money* as the correct translation. A second type of error involved translating the distractor instead of the target word. For

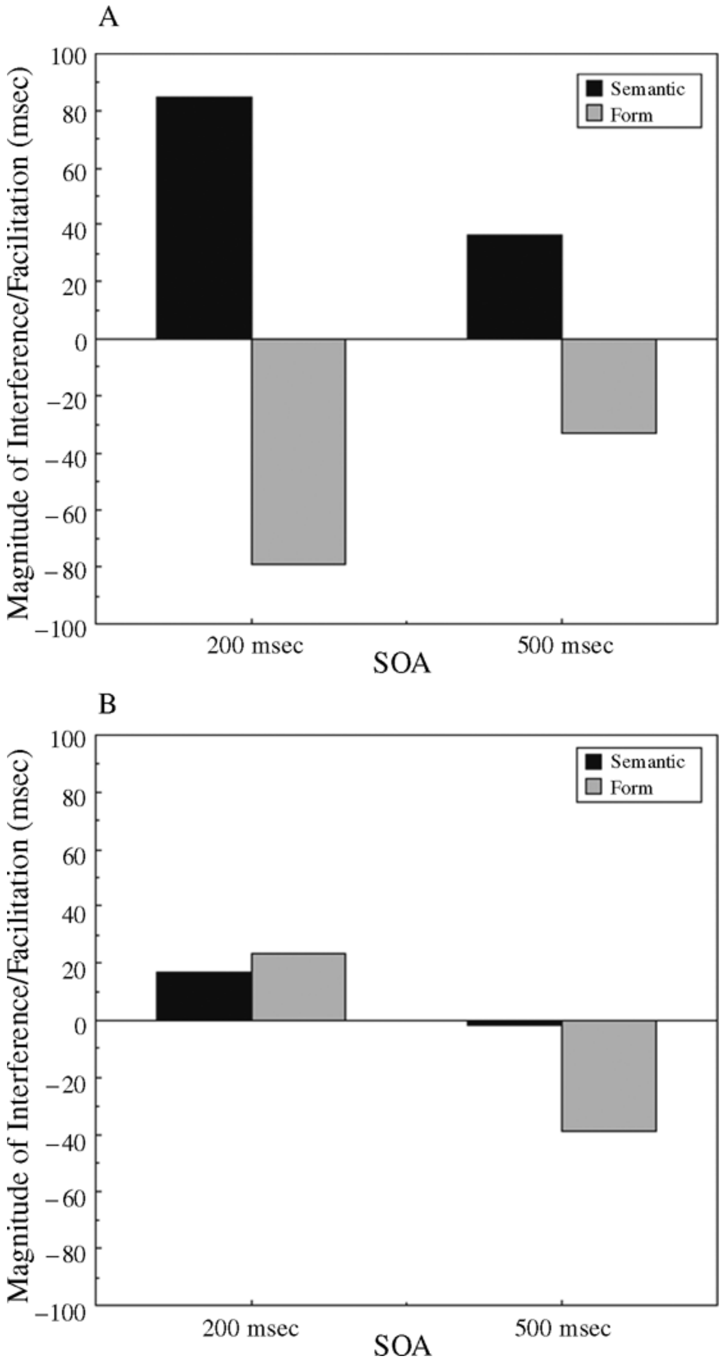

Figure 2. The magnitude of facilitation (negative) and interference (positive), in milliseconds, for semantically related distractors relative to controls and form related distractors relative to controls as a function of stimulus onset asynchrony (SOA, 200 vs. 500 msec) when (A) the distractors appeared in the language of output in Experiment 1 and (B) the distractors appeared in the language of input in Experiment 2.

example, if the word *vestido* was presented for translation and the distractor was *press*, we asked whether the participants ever incorrectly said *apretón*, rather than *dress*, as the correct translation.

To compare the results of the two experiments, separate ANOVAs were performed on the percentages of errors that occurred when the participants made the mistake of naming the distractor or the mistake of translating the distractor. In these analyses, experiment and SOA were between-group factors, whereas direction of translation, relatedness of the distractor, and type of distractor were within-group factors. The percentages of errors that occurred when the participants named the distractor are shown in Table 5. The main finding was that this type of mistake was much more likely to occur in Experiment 1 than in Experiment 2 and that, within Experiment 1, it was more likely to occur when the distractor was presented at the short-SOA condition and was semantically related. The ANOVA revealed a signficant interaction between SOA, relatedness of distractor, and type of distractor [$F_1(1,59) = 8.99$, $MS_e = 0.008$, $p < .01$], providing statistical support for the claim that the participants were more likely to incorrectly produce the distractor itself in the short-SOA condition when the distractor was semantically related. A significant interaction was also obtained between experiment and direction of translation [$F_1(1,59) = 4.05$, $MS_e = 0.017$, $p < .05$]. A simple effects test revealed a significant effect of experiment for L2 to L1 translation [$F_1(1,118) = 8.88$, $MS_e = 0.179$, $p < .01$]. In both experiments, the participants were more likely to name the distractor words when they appeared in L1. However, the largest percentage of errors of this type occurred in Experiment 1 in the L2 to L1 direction in which an L1 distractor matched the language of production required by the task.

Table 5
Percentages of Errors in Which Participants Named or Translated the Distractor in Both Experiments 1 and 2

	Distractor Relatedness			
	Experiment 1		Experiment 2	
Condition	Related	Unrelated	Related	Unrelated
	L1–L2			
Short SOA				
Form				
Name	3.1	0.0	0.0	0.0
Translate	0.0	0.0	0.0	0.0
Semantic				
Name	10.9	0.0	6.2	0.0
Translate	0.0	0.0	6.2	0.0
Long SOA				
Form				
Name	1.6	0.0	5.5	0.0
Translate	0.0	0.0	2.2	0.0
Semantic				
Name	0.0	0.0	2.2	0.0
Translate	0.0	0.0	8.9	0.0
	L2–L1			
Short SOA				
Form				
Name	3.1	6.2	0.0	0.0
Translate	0.0	0.0	15.6	0.0
Semantic				
Name	16.9	2.1	0.0	0.0
Translate	0.0	0.0	6.2	2.1
Long SOA				
Form				
Name	6.2	0.0	1.3	0.0
Translate	0.0	0.0	4.7	4.0
Semantic				
Name	5.6	0.0	0.0	0.0
Translate	0.0	0.0	5.0	4.4

Note—The data are given as a function of the direction of translation, SOA, type of distractor (form or semantic), and the relatedness of the distractor.

A marginal interaction between relatedness of distractor and type of distractor was also observed [$F_1(1,59) = 3.76$, $MS_e = 0.008$, $p < .06$]. This interaction indicates that more errors occurred for distractors that were semantically related than for distractors that were related by word form or were unrelated. There were also main effects of relatedness of the distractor [$F_1(1,59) = 7.22$, $MS_e = 0.020$, $p < .01$] and of experiment [$F_1(1,59) = 4.92$, $MS_e = 0.016$, $p < .05$], with more errors occurring in Experiment 1 than in Experiment 2. The results suggest that production is more likely to result in an error when the distractor is semantically related and in the language of production. The error data thus converge with the translation RTs in suggesting that only a small set of lexical candidates are active prior to production. If the distractor itself is among the competing responses, the probability that the distractor itself will be spoken is high, relative to other conditions.

The percentages of errors that occurred when the participants translated the distractor are also shown in Table 5. The main finding was that the mistake of translating the distractor occurred only in Experiment 2. In Experiment 1, in which the distractor appeared in the language of production, the participants never made the error of incorrectly translating the distractor. Furthermore, the striking result is that a relatively high percentage of these mistakes occurred when the participants translated from L2 to L1 and the distractors were form-related L2 words. Although the effect of form distractors was not significant in the overall analysis of RTs in Experiment 2, if we look at the RT data in Table 4, we see that form interference in Experiment 2 occurred when form-related distractor words were presented during L2 to L1 translation at the short SOA. The error data thus converge with the latency data to suggest that there was increased lexical competition in recognizing the L2 word induced by the presence of a second L2 word that was orthographically and/or phonologically related to the target word.

To examine the pattern of errors more closely, an ANOVA was then performed on these data. The analysis revealed a significant interaction between experiment and relatedness of the distractor [$F_1(1,29) = 12.68$, $MS_e = 0.006$, $p < .01$]. A main effect of relatedness of distractor was also observed [$F_1(1,59) = 12.68$, $MS_e = 0.006$, $p < .01$]. More errors occurred when the distractor was related to the target than when it was unrelated to the target. The main effect of experiment was also significant [$F_1(1,59) = 9.73$, $MS_e = 0.018$, $p < .01$], as was expected, given that no errors of this type occurred in Experiment 1.

The analysis of production errors suggests that the language of production is selected early in processing during translation, because the error of incorrectly naming the distractor was more likely to occur when a semantically related distractor word in the language of production was presented at a short SOA. These results suggest that when the distractor is among the competing responses, it becomes a viable candidate for output, and the error of naming the distractor is more likely to occur.

Implications for Models of Bilingual Production

The results of the two experiments reported in this paper suggest that, contrary to the evidence based on cross-language picture–word interference (e.g., Costa et al., 1999; Hermans et al., 1998), language selection occurs relatively early in translation. In the picture–word task, distractors in either of the bilingual's two languages produce semantic interference regardless of their match to the language of production. In translation, these effects are observed only when the language of the distractor matches the language of production, suggesting that during the process of lexicalization, only competitors from the language to be spoken are active.

Why might language selection differ for picture naming and translation? We hypothesize that translation provides a language cue that gives the bilingual specific information about the language and the word that is not to be produced. Picture naming does not provide a language cue unless the picture conveys culturally or linguistically specific information. In most picture-naming studies, the line drawings that are typically used as stimuli do not provide cues that bias the selection of one language rather than the other. Furthermore, in translation, the bottom-up processes associated with the recognition of the target word stimulus will activate a set of lexical form candidates (see, e.g., Andrews, 1997; Van Heuven et al., 1998). However, rarely will it be the case that these lexical neighbors are also semantic relatives. In contrast, when a picture is recognized, a cohort of related visual representations will be activated, and the structurally similar objects are also likely to be semantically related (e.g., Vitkovitch, Humphreys, & Lloyd-Jones, 1993). Thus, in picture naming there are at least two factors working against early language selection: There is no language cue in the stimulus itself, and the process of resolving the identity of the object is likely to have the consequence of causing the activation of semantic alternatives that may, in turn, activate their corresponding lexical representations in both languages. If the information about the language of production can be encoded in the translation by virtue of the presence of the language cue contained within the input itself, activation of nontarget competitors may be effectively inhibited earlier in processing than it is in the comparable picture-naming task, in which no language cues are present.

Figure 3 shows two production models, one for picture naming (Figure 3A) and the other for translation (Figure 3B), adapted from the models proposed by Poulisse and Bongaerts (1994) and Hermans (2000). Each model includes three levels: a conceptual level, in which there are conceptual cues and a language cue, a lemma level, consisting of abstract lexical representations for words in each language, and a phonological level, in which the phonological features of to-be-spoken words are represented. The conceptual cues represent the semantic features of the target, and the language cue marks the language that the bilingual intends to speak. The model for picture naming assumes nonselective access for both languages so that having the intention to speak in one language only does not effectively eliminate activity among lemmas from the nontarget language. To the contrary, according to the model shown in Figure 3A, lemmas in both of the bilingual's languages are active prior to naming the object. However, as Hermans et al. (1998) and Costa et al. (1999) have argued, language selection appears to occur in picture naming at the lemma level, so that the phonological features of the translation equivalent (in this case for the word *silla* in Spanish) would not be active.

In Figure 3B, we modified the picture-naming model to represent one of the ways in which we believe that translation differs from picture naming—namely, to provide a strong language cue to bias selection toward the intended language of production. Here, the target word to be translated (*silla*) activates not only its respective meaning, but also the language cue. The presence of the target word in Spanish provides information in the context of the task schema for translation (Green, 1998) to enable the language cue to be activated directly. The joint activation of the conceptual cues and the language cue will serve to bias the activation of English lemmas over Spanish lemmas, allowing language selection to occur prior to the lemma level. When distractors are subsequently presented in the language of the spoken output, normal Stroop-type interference will be observed. However, when distractors are presented in the language of the target input, they will no longer produce interference, because they are not competing for selection.

For the sake of simplicity, we have refrained from adding additional connections in the translation model that would depict the consequences of the bottom-up processes that must be engaged when the target word is recognized. Likewise, the picture-naming model does not include an additional level that would represent the computation of a structural description for the object and the corresponding activation of related visually similar representations. However, as was described above, we believe that these bottom-up processes function to further differentiate the processing engaged by the two production tasks. In the case of translation, the activation of lexical neighbors that resemble the target in form, but not in meaning, may serve to inhibit semantic alternatives in the target language that are weakly activated by top-down processes. That is, this process may further inhibit the influence of semantic competitors in the target language, relative to those in the language to be produced. The corresponding bottom-up processes in picture naming will have just the opposite effect. Visually similar competitors, unlike word neighbors, will

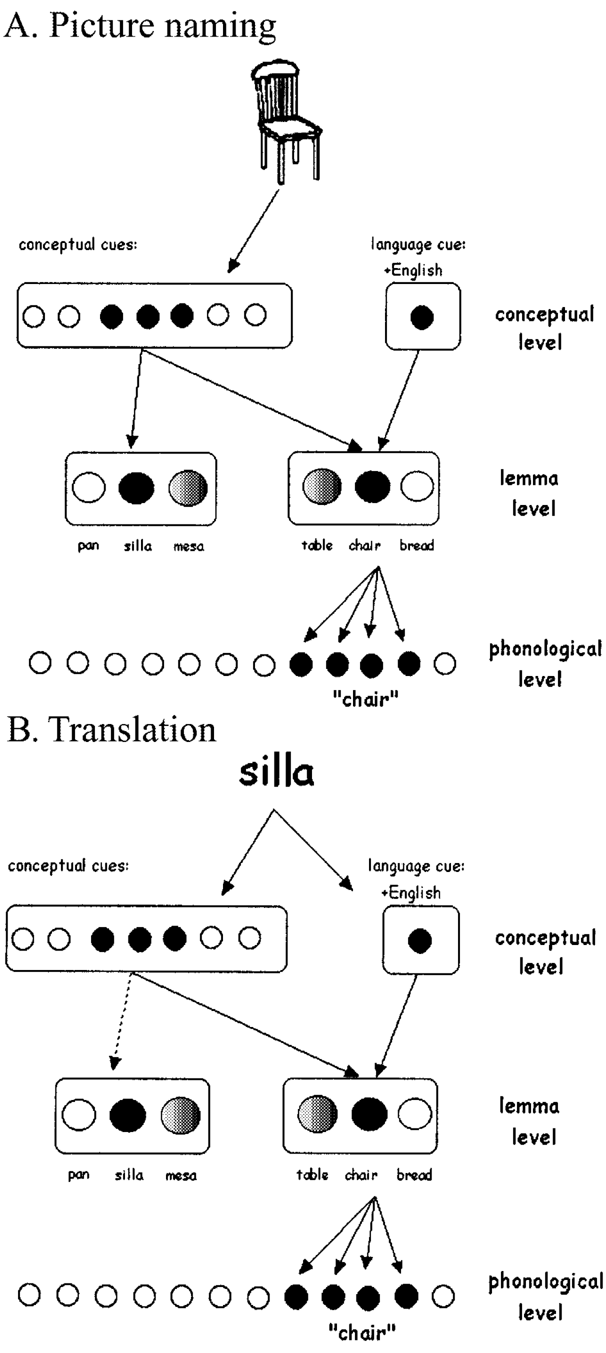

Figure 3. Models of bilingual language production for (A) picture naming and (B) translation, adapted from Poulisse and Bongaerts (1994) and Hermans (2000).

also tend to activate semantically related representations. One consequence may be to extend the time course of resolving the identity of the object. In the absence of an additional cue, the effect of an extended time course during the initial stages of object recognition will be to activate corresponding lexical alternatives in both languages.

Picture Naming Versus Translation

A number of past studies have compared picture naming and translation as a basis on which to infer the nature of the connections between words and concepts in bilingual memory (e.g., Chen, Cheung, & Lau, 1997; Potter et al., 1984; Sholl et al., 1995). In virtually every one of these studies, the difference between the time to produce words in L1 versus L2 is larger in picture naming than in translation. One interpretation of the language difference in picture naming is that, especially for naming in L2, there is an additional processing demand of resolving competition from L1 candidates.

A second source of evidence for the difference between picture naming and translation was reported by Sholl (1996), in a study that examined the effects of conceptual animacy in these tasks. Conceptual animacy (McRae, de Sa, & Seidenberg, 1997) is highly correlated with the degree of visual similarity among semantically related items. Objects drawn from biological categories tend to be more visually similar than those from nonbiological categories. Sholl used the comparison between picture naming and bilingual translation to tease apart the contributions of visual similarity and conceptual animacy. She found inhibitory effects for animate objects in picture naming but facilitation for the same concepts in translation, again suggesting that two tasks share component processes only to a degree and that picture naming is likely to reflect an initial process of resolving competition among visually and semantically related alternatives.

CONCLUSIONS

In two translation-Stroop experiments, we demonstrated that semantic interference and form facilitation are observed reliably only when the language of the distractor word matches the language to be spoken. Unlike picture–word interference studies, in which these effects are observed regardless of the match between the language of the distractor and the language of production, the translation results suggest that the locus of language selection depends on available cues. When cues are present in the form of the language input, the bilingual appears to be able to selectively activate lexical alternatives in the intended language of speaking. When those cues are absent, lexical candidates appear to be activated in both languages and subsequently compete for selection.

To what extent do the results of experiments on bilingual production out of context tell us something general about spoken language? Presumably, most of the consequences of parallel activation of lexical candidates in both languages are resolved in context, where bilingual speakers do not often mix words unless they are intentionally code switching with another bilingual. What is of interest is that we know very little about the cues that allow bilinguals to control their performance. The results of the present study suggest that the form of the input that initiates the production process may play an important role in determining the locus of language selection. It will be critical in future work to consider how these mechanisms operate in and out of context.

REFERENCES

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review*, **4**, 439-461.

Chen, H.-C., Cheung, H., & Lau, S. (1997). Examining and reexamining the structure of Chinese–English bilingual memory. *Psychological Research*, **60**, 270-283.

Chen, H.-C., & Leung, Y.-S. (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 316-325.

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 & Language*, **41**, 365-397.

Francis, W. N., & Kučera, H. (1982). *Frequency analysis of English usage*. Providence, RI: Brown University Press.

Glaser, W. R., & Düngelhoff, F.-J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 640-654.

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

Hermans, D. (2000). *Word production in a foreign language*. Unpublished doctoral dissertation, University of Nijmegen.

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 & Cognition*, **1**, 213-230.

Jescheniak, J. D., & Schriefers, H. (1998). Discrete serial versus cascaded processing in lexical access in speech production: Further evidence from the coactivation of near-synonyms. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1256-1274.

Kroll, J. F., & Curley, J. (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second language words. In M. Gruneberg, P. Morris, & R. Sykes (Eds.), *Practical aspects of memory* (Vol. 2, pp. 389-395). London: Wiley.

Kroll, J. F., & de Groot, A. M. B. (1997). Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 169-199). Mahwah, NJ: Erlbaum.

Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory & Language*, **33**, 149-174.

La Heij, W., de Bruyn, E., Elens, E., Hartsuiker, R., & Helaha, D. (1990). Orthographic facilitation and categorical interference in a word-translation variant of the Stroop task. *Canadian Journal of Psychology*, **44**, 76-83.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral & Brain Sciences*, **22**, 1-75.

Lloyd-Jones, T. J., & Humphreys, G. W. (1997). Categorizing chairs and naming pears: Category differences in object processing as a function of task and priming. *Memory & Cognition*, **25**, 606-624.

Lupker, S. J. (1979). The semantic nature of response competition in the picture–word interference task. *Memory & Cognition*, **7**, 485-495.

McRae, J., de Sa, V. R., & Seidenberg, M. S. (1997). On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, **126**, 99-130.

Peterson, R. R., & Savoy, P. (1998). Lexical selection and phonological encoding during language production: Evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 539-557.

Potter, M. C., So, K.-F., von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning & 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.

Sánchez-Casas, R. M., Davis, C. W., & García-Albea, J. E. (1992). Bilingual lexical processing: Exploring the cognate–noncognate distinction. *European Journal of Cognitive Psychology*, **4**, 293-310.

Sholl, A. (1996). *Animacy effects in picture naming and bilingual translation: Perceptual and semantic contributions to concept-mediation*. Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Sholl, A., Sankaranarayanan, A., & Kroll, J. F. (1995). Transfer between picture naming and translation: A test of asymmetries in bilingual memory. *Psychological Science*, **6**, 45-49.

Starreveld, P. A. (2000). On the interpretation of onsets of auditory context effects in word production. *Journal of Memory & Language*, **42**, 497-525.

Starreveld, P. A., & La Heij, W. (1995). Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 686-698.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.

Talamas, A., Kroll, J. F., & Dufour, R. (1999). From form to meaning: Stages in the acquisition of second language vocabulary. *Bilingualism: Language & Cognition*, **2**, 45-58.

Taylor, T. E., & Lupker, S. J. (2001). Sequential effects in naming: A time-criterion account. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 117-138.

Van Heuven, W. J. B., Dijkstra, A., & Grainger, J. (1998). Orthographic neighborhood effects in bilingual word recognition. *Journal of Memory & Language*, **39**, 458-483.

Vitkovitch, M., Humphreys, G. W., & Lloyd-Jones, T. J. (1993). On naming a giraffe a zebra: Picture naming errors across different object categories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 243-259.

NOTES

1. We use the term *bilingual* to refer to anyone who uses an L2 at a relatively high level of proficiency. Because most of the research we will describe concerns individuals who are late L2 learners, few are balanced bilinguals.

2. An alternative interpretation of the translation asymmetry is given by the inhibitory control model proposed by Green (1998). By this account, L1 to L2 translation requires the suppression of L1, whereas L2 to L1 translation does not require comparable suppression of L2, because L2 is hypothesized to be less active. This may also provide an explanation for the elimination of the asymmetry in Experiment 2. During L2 to L1 translation, both languages are active; however, the presentation of an L2 distractor may have the effect of increasing the activation of L2. The increase in activation of L2 may therefore disrupt the production process because L2 must be suppressed before L1 can be produced. The consequence of increased activation of L2 in the L2 to L1 translation direction will be to increase the inhibitory control demands so that L2 to L1 will resemble L1 to L2. For translation from L1 to L2, the L1 distractor will not produce a corresponding increase in L1 activation, because L1 is already at a relatively high level of activation. Lexical competition would be expected to increase when the distractor is in the language of the target. However, L2 to L1 translation should be more affected than L1 to L2 translation, because L2 will be more vulnerable to the consequences of within-language lexical competition (see Talamas, Kroll, & Dufour, 1999).

3. A series of *t* tests was also performed on the data of Experiment 2 to be certain that none of the differences was significant. In each of these comparisons, performed separately for each distractor type at each SOA and for each direction of translation, one-tailed tests revealed *p* values greater than .10, confirming the pattern revealed by the omnibus ANOVA.

(Manuscript received March 15, 2001;
revision accepted for publication January 10, 2002.)