When do false memories cross language boundaries in English-Spanish bilinguals?

BROOKE H. SAHLIN, MATTHEW G. HARDING, and JOHN G. SEAMON
Wesleyan University, Middletown, Connecticut

To determine whether false memories cross language boundaries, we presented English-Spanish bilinguals with conceptually related word lists for five study-test trials. Some lists were heard in English, some in Spanish, and they were then followed by a recognition memory test composed of studied words, conceptually related nonstudied critical words, and unrelated words presented in the same language used at study or in a different language. Even though participants were instructed to recognize only previously heard words, they falsely recognized both same-language and different-language critical words. With practice, participants increased their accurate recognition of list words and decreased their false recognition of critical words when the study-test language was the same, and they decreased their false recognition of list words and critical words when the language differed. False memories can cross language boundaries when participants rely on conceptual representations from the word lists, but these errors decrease over trials as participants increasingly rely on language-specific lexical representations.

Over the past decade, there has been an explosion of research on false memory. False memory is demonstrated when a person attributes erroneous details to an actual event or remembers an event that never took place. Although much of this research has involved memory for linguistic stimuli (see Roediger & Gallo, 2002, for a review), surprisingly few of these studies have involved bilingual participants for whom language was varied between study and test. This shortcoming is especially noteworthy because at least half of the world's population is thought to be bilingual (French & Jacquet, 2004). To date, only two studies, involving different procedures, have addressed the issue of whether false memories can cross language boundaries.

In the postevent misinformation procedure, witnesses are exposed to an event, after which they receive inaccurate information about the event. Participants often incorporate this information into their subsequent recollections (see, e.g., Loftus, Donders, Hoffman, & Schooler, 1989). Shaw, Garcia, and Robles (1997) found that bilingual witnesses were equally susceptible to postevent misinformation that was presented in the same language as the original event or in a different language. This study suggested that false memories can cross language boundaries.

In the Deese/Roediger-McDermott (DRM) procedure, people study lists composed of conceptually related words (e.g., *door, glass, pane*). Later, they often recall or recognize a nonstudied converging associate (e.g., *win-*

dow), known as a *critical word*, as frequently as studied words. Previous DRM studies have demonstrated a strong false-memory effect using monolingual word lists in English (e.g., Deese, 1959; Roediger & McDermott, 1995; Seamon et al., in press), Spanish (e.g., García-Bajos & Migueles, 1997; Pérez-Mata, Read, & Diges, 2002), or Portuguese (Stein & Pergher, 2001), but only one study has tested bilingual participants. Kawasaki-Miyaji, Inoue, and Yama (2003) gave participants a multiple-choice recognition test in which some lists were tested in the same language as at study (e.g., Japanese/Japanese) and others in a different language (e.g., Japanese/English). Although more false-recognition errors occurred for words in the Japanese language, in which participants were more proficient, the participants falsely recognized nonstudied critical words in both the same-language and different-language conditions. This study also suggested that false memories can cross language boundaries after a single study trial.

The present research adds to the sparse literature on bilingual false memory by specifying, using the DRM procedure, the conditions under which same-language and different-language false memories occur in proficient bilingual participants. We conducted our study without prior knowledge of Kawasaki-Miyaji et al.'s (2003) research, and the two studies differ in important ways. First, the targeted languages used in these studies were not the same. Second, we employed a multiple- rather than a single-trial study-test format in order to observe how same-language and different-language false memories could change over trials. Third, we presented the study lists auditorily, not visually. Finally, we tested participants who were highly proficient in two languages as a result of sustained exposure to English and Spanish from infancy through young adulthood.

The authors express their appreciation to Catherine Best for use of her audio recording equipment and sound-controlled test rooms. Correspondence should be addressed to J. G. Seamon, Department of Psychology, Wesleyan University, Middletown, CT 06459-0408 (e-mail: jseamon@wesleyan.edu).

Briefly, we presented DRM word lists auditorily at study, some lists in English, others in Spanish. Following study, a yes–no recognition memory test composed of studied words, conceptually related critical words, and unrelated words was presented in either the same language used at study (English/English or Spanish/Spanish) or a different language (Spanish/English or English/Spanish). The same procedure was employed for five successive study–test trials, and the participants were instructed to recognize only previously heard words on each trial.

With monolingual DRM lists, previous research has shown that accurate recognition increases and false recognition for critical words decreases over trials (Kensinger & Schacter, 1999; Seamon et al., 2003). Participants encode the themes of the lists quickly, leading to false memory for critical words, but they gradually learn which words were in the studied lists, and which words, although conceptually related, were not, leading to a decrease in false memory over trials. With bilingual lists presented over five study–test trials, the same results should be obtained from proficient bilingual participants for lists that are studied and tested in the same language.

For lists that are studied and tested in different languages, and when participants are instructed to recognize only those test items in the same language used at study, different predictions are possible depending upon the underlying model of bilingual representation. If word information is represented in language-specific lexicons, little false recognition for critical words should occur when the language changes between study and test. False recognition of critical words should occur only when the language is the same for study and test, and these false recognitions should decrease over trials.

However, many theorists have proposed that bilinguals can have separate lexical representations for a word in each language, but a common conceptual representation for the word in both languages (e.g., French & Jacques, 2004; Kroll & de Groot, 1997; Smith, 1997). The processing demands of a task have also been shown to influence whether studies show cross-language or only same-language effects (Durgunoğlu & Roediger, 1987; Gollan & Kroll, 2001; Zeelenberg & Pecher, 2003). For example, semantic priming, picture naming, and paired-associate studies have observed cross-language facilitation and/or interference effects with bilinguals (Chen & Ng, 1989; Fox, 1996; Keatley, Spinks, & de Gelder, 1994; MacLeod, 1976), suggesting that participants can access language-independent conceptual representations when processing meaning from either of their two languages (see Smith, 1997, for a review). On the other hand, studies involving repetition priming, lexical decision, and word fragment completion typically show same-language but not cross-language facilitation for translation equivalents (Durgunoğlu & Roediger, 1987; Gerard & Scarborough, 1989; Watkins & Peynircioğlu, 1983). These tasks typically draw on lexical processing and thus reflect language-specific lexical representations.

Because the processing requirements of a task determine whether cross-language effects are observed and be-

cause false memories in the DRM task reflect conceptual, not lexical, processing errors (the critical words are never presented), false recognition of critical words should be evident in both the same-language and different-language conditions in the present study, at least during initial learning. A different-language critical word is conceptually equivalent to a same-language critical word but lexically different. Thus, both same-language and different-language critical words should produce false memory during the initial trials. Over trials, though, participants should reduce their false memories of critical words because they learn that conceptual relatedness to the studied words is necessary but not sufficient for recognition; a test word must also match a studied word in terms of its lexical representation. Consequently, over trials, bilingual participants should reduce their false recognitions not only of list words presented in a different language, but also of critical words presented in either the same or a different language. These results would demonstrate both same-language and cross-language false memories in the DRM task. In addition, as list learning proceeds over multiple study–test trials, both types of false memories should diminish as participants edit their memory by increasingly relying on lexical representations for recognition decisions. This study tested these predictions.

METHOD

Participants

The participants were 20 English–Spanish bilinguals, 9 females and 11 males, between the ages of 18 and 27 years ($M = 20$, $SD = 1.96$). They were Wesleyan University undergraduates who served as paid volunteers, and none had taken part in any related memory research. All participants were qualified as “highly proficient bilinguals” based on information they provided on an extensive language background questionnaire. This questionnaire, which elicited information on language exposure settings, length of exposure, and age at time of exposure, also required the participants to rate their abilities to speak, understand, read, and write English and Spanish on a 5-point scale (with 5 the highest level). The self-report ratings were high for both languages, although proficiency was marginally higher for English than for Spanish. For English, the ratings (expressed as means and standard deviations) were: for speaking (5, 0), for understanding (5, 0), for reading (5, 0), and for writing (5, 0). For Spanish, the ratings were: for speaking (4.58, 0.49), for understanding (4.78, 0.41), for reading (4.50, 0.69), and for writing (4.35, 0.80). All participants had at least one native Spanish-speaking parent at home, and the mean ages at which participants began learning both languages was 0.90 years ($SD = 2.22$) for English and 0.84 years ($SD = 2.54$) for Spanish.

Materials

We used Stadler, Roediger, and McDermott’s (1999) false-memory norms to select 24 DRM lists that produced high levels of false recognition in English. These lists were then translated into Spanish, although a one-to-one translation of a word was not always available between languages. For example, some English lists contained multiple words for which there was a single word in Spanish (e.g., *thief*, *robber*, and *burglar* are represented in Spanish by *ladron*). Similarly, there were also cases in which only one English word covered the meanings for multiple Spanish words. In addition, some words were not used because their membership in a list was based on an idiomatic association that was culturally constrained or language specific (e.g., the *needle–haystack* association does not exist in

Spanish). Finally, homographs that were spelled the same in English and Spanish were eliminated to avoid confusion on the recognition tests. To overcome these difficulties, we selected the first 10 words in each 15-word DRM list that enabled us to translate the words accurately from English to Spanish (see the Appendix). A native Spanish-speaking informant verified all translations.

From the Stadler et al. (1999) norms, the 24 lists were divided into two sets of 12 lists (labeled Sets A and B) that were comparable for false recognition. Half of the participants received Set A for study, and half received Set B. The set not used during study provided distractors for the recognition test. Each list was composed of 10 words that were all converging associates of the nonstudied critical word. The strongest associates to the critical word normally occurred first in each list, and the order of the words was constant over all conditions for a participant. Of the 12 lists in each set, 6 were presented in English and 6 in Spanish, in an alternating pattern that began with English for half of the participants and Spanish for the other half.

We created an audio CD of each list set and language order by employing a highly proficient bilingual woman who recorded the word lists in English and Spanish. Each CD track contained a total of 120 words presented at a rate of about one word every 3 sec, with a tone separating each blocked list. Digital recordings (44.1 kHz) were made directly to an Apple iBook hard drive with Praat speech recording and editing software using a Nakamichi CM-100 electret condenser microphone (range: 30–1800 Hz) powered by a Nakamichi MX-100 microphone preamplifier. The visual word recognition tests were presented on Apple iBooks using PowerLab software.

Design

The two between-participants variables were list set (Set A vs. Set B) and language alternation pattern within a list set (English first vs. Spanish first). The three within-participants variables were language at study and test (same vs. different language), study–test trial (1–5), and word type at test (studied vs. nonstudied list words, and related vs. unrelated critical words).

Procedure

At the start of the experiment, the participants, who were tested individually in sound-controlled rooms, were told that the purpose of the experiment was to test their memory for orally presented word lists. The participants were informed that they would hear 12 lists of 10 words, some lists in English and some in Spanish. The participants were told to listen carefully to the word lists at study so they would remember as many words as possible for the subsequent recognition test. The presentation of the word lists was followed by a visual recognition memory test that contained both studied and nonstudied words. Other than noting that some test words were studied and others were not, the participants were not informed about the composition of the test words or the false-memory effect. The participants were told that this same procedure, using the same word lists, would be employed on five consecutive study–test trials.

For each of the five trials, the recognition test was administered immediately after the last study list was heard. The recognition test contained 96 words presented on a computer screen one at a time, with the words YES and NO below each word. For each test word, the participants were told to indicate whether they had heard “that word in that language” during study by clicking either YES or NO with the mouse. A test word remained on the screen until a response was made. For each of the five study–test trials, the recognition test consisted of the same 96 words presented in a different random order on each trial. The test consisted of three items from each studied list (Serial Positions 1, 8, and 10), the nonstudied critical word from each studied list, three items from each nonstudied list (Serial Positions 1, 8, and 10), and the critical word from each nonstudied list. Three lists studied in English and three lists studied in Spanish were tested using words from those lists in the same language as at study; for the other three English and three Spanish lists at study, the lan-

guage was switched at test. Following the last trial, the participants were debriefed.

RESULTS

The results are presented in terms of correct and false recognition, as well as of sensitivity by the high-threshold measure. Previous research has shown that the high-threshold measure tends to be more sensitive to change than the signal detection or the nonparametric measures (Seamon, Luo, et al., 2002; Snodgrass & Corwin, 1988). The estimate of sensitivity, Pr , is defined as hits minus false alarms, with higher scores indicating greater sensitivity than do lower scores. Two sets of Pr measures were computed to provide separate estimates involving item-specific memory and false-memory susceptibility. For item-specific memory, the mean estimates of sensitivity were computed from hits and false alarms to studied and nonstudied list words from individual participants' responses. For false-memory susceptibility, these estimates were computed from false alarms to related and unrelated critical words. This procedure, used previously by researchers (e.g., Kensinger & Schacter, 1999; Seamon, Lee, et al., 2002), treated the false recognition of related critical words as “hits” to provide a measure of the extent to which participants were fooled by these nonstudied words.

Mean Recognition Proportions

The mean recognition proportions for each language condition, study–test trial, and word type are presented in Table 1. Because list set (A vs. B) and language alternation pattern (English first vs. Spanish first) did not systematically affect performance in any analysis, they will not be considered further. Table 1 shows that when the study and test languages were the same, the mean proportions of recognized words were greater for studied than for nonstudied list words and for related than for unrelated critical words for both languages and all trials. These results demonstrate accurate memory for list words and false memory for related critical words. Consistent with previous research involving multiple study–test trials in the same language (Kensinger & Schacter, 1999; Seamon et al., 2003), accurate recognition of list words increased over the five trials, whereas false recognition of critical words decreased. Table 1 also shows that when the study and test languages were different, participants also recognized (incorrectly) more studied than nonstudied list words and more related than unrelated critical words. These erroneous responses due to language decreased over trials for both list words and critical words as participants gradually learned that accurate recognition of a test item required both the meaning and language of the item to be the same as during study. Most importantly, Table 1 clearly indicates that false recognition of critical words exceeded that of list words presented in a different language, especially in the initial trials, suggesting that false memories can cross language boundaries.

Table 1
Mean Recognition Proportions for List Words and
Critical Words in the Same or in a Different Language

Word Condition	Study-Test Trials					<i>M</i>
	1	2	3	4	5	
English-English						
List words						
Studied	.62	.86	.91	.94	.93	.85
Nonstudied	.08	.05	.03	.02	.01	.04
Critical words						
Related	.75	.72	.70	.60	.53	.66
Unrelated	.11	.08	.04	.03	.03	.06
Spanish-Spanish						
List words						
Studied	.69	.86	.93	.95	.96	.88
Nonstudied	.11	.07	.06	.03	.03	.06
Critical words						
Related	.62	.50	.52	.37	.33	.47
Unrelated	.08	.08	.03	.01	.00	.04
English-Spanish						
List words						
Studied	.22	.25	.16	.13	.09	.17
Nonstudied	.11	.07	.06	.03	.03	.06
Critical words						
Related	.38	.28	.27	.17	.17	.25
Unrelated	.08	.08	.03	.01	.00	.04
Spanish-English						
List words						
Studied	.20	.12	.12	.11	.10	.13
Nonstudied	.08	.05	.03	.02	.01	.04
Critical words						
Related	.33	.23	.20	.10	.08	.19
Unrelated	.11	.08	.04	.03	.03	.06

Note—*M* represents the mean proportion across trials.

True-Memory Sensitivity and False-Memory Susceptibility

Because of differences in the baseline false alarm rates over language conditions and trials, the data were analyzed in terms of sensitivity by the high-threshold measure. These data, shown in Table 2, reveal the same pattern of results observed in Table 1. Generally similar results were obtained for English and Spanish from our proficient bilingual participants. When the words were studied and tested in the same language, accurate recognition of list words increased and false recognition of critical words decreased over trials. When the words were studied and tested in a different language, false recognition of list words and critical words decreased over trials, and the false recognition of critical words was greater than the false recognition of list words for the initial trials. The latter finding, together with the finding that false recognition of critical words was less in a different language than in the same language (see the means in Table 2), strongly implies that although false memories can cross language boundaries, they are not language free. Participants encode not only the meaning of the words presented at study, but also the language of the words as they are presented, and participants increasingly use language to aid their recognition decisions as list learning proceeds over trials.

These observations were supported by the results of analyses of variance. The first set of analyses of the *Pr* measure was conducted on the same-language conditions, and it included language (English/English vs. Spanish/Spanish), word type (list words vs. critical words), and trials (1–5) as within-participants variables. Main effects of language [$F(1,18) = 5.75$, $MS_e = 0.72$, $p < .05$], word type [$F(1,18) = 77.77$, $MS_e = 9.07$, $p < .001$], and trials [$F(4,72) = 5.83$, $MS_e = 0.24$, $p < .001$] were found, as well as interactions of language and word type [$F(1,18) = 12.24$, $MS_e = 0.84$, $p < .01$] and trials and word type [$F(4,72) = 29.86$, $MS_e = 0.89$, $p < .001$]. As shown by the means for the same-language conditions in Table 2, accurate recognition was the same for English (.81) and for Spanish (.82) ($F < 1$), but false recognition was greater for English (.60) than for Spanish (.43) [$F(1,18) = 9.04$, $MS_e = 0.17$, $p < .01$]. The false-recognition level for English words was comparable to that observed in previous research with only English words over multiple study-test trials (see, e.g., Kensinger & Schacter, 1999; Seamon et al., 2003), but the false-recognition level of Spanish words, though substantial, was lower than that for English words. This language difference in false recognition may reflect a subtle proficiency difference in our bilingual participants, as suggested by the results of our self-report questionnaire. Kroll and de Groot (1997) have noted that even when bilinguals are highly proficient in both languages, performance asymmetries are frequently observed.

A second set of analyses, employing the same variables, was conducted on the different-language conditions. There was no main effect of language (English/Spanish vs. English/Spanish; $p > .15$), but main effects of word type [$F(1,18) = 13.19$, $MS_e = 0.53$, $p < .01$] and trials [$F(4,72) = 7.72$, $MS_e = 0.18$, $p < .001$] did emerge, as well as an interaction of trials and word type [$F(4,72) = 3.63$, $MS_e = 0.06$, $p < .01$]. No other effects were reliable. After collapsing over the two different-language condi-

Table 2
Mean *Pr* Sensitivity Scores for List Words and
Critical Words in the Same or in a Different Language

	Study-Test Trials					
Word Condition	1	2	3	4	5	<i>M</i>
English-English						
List words	.54	.80	.87	.92	.92	.81
Critical words	.64	.64	.66	.57	.50	.60
Spanish-Spanish						
List words	.58	.79	.87	.92	.93	.82
Critical words	.54	.43	.49	.35	.33	.43
English-Spanish						
List words	.13	.20	.12	.11	.07	.12
Critical words	.32	.22	.24	.15	.16	.22
Spanish-English						
List words	.14	.08	.11	.09	.09	.10
Critical words	.25	.18	.19	.08	.06	.15

Note—*Pr*, hits minus false alarms; *M*, mean *Pr* across trials.

tions in Table 2, the results show that false recognition of critical words was greater than false recognition of list words in a different language for Trials 1 [.29 vs. .14; $F(1,18) = 14.82$, $MS_e = 0.03$, $p < .001$] and 3 [.21 vs. .11; $F(1,18) = 15.04$, $MS_e = 0.01$, $p < .001$]. False recognition was numerically greater for critical words than for list words on Trials 2 (.20 vs. .14), 4 (.12 vs. .10), and 5 (.11 vs. .08), but these differences were not reliable (all $ps > .15$).

Finally, to determine whether false recognition of critical words was greater when study–test languages were the same than when they were different, we conducted two additional analyses. In the first, we compared the English/English condition with the English/Spanish condition; in the second, we compared the Spanish/Spanish condition with the Spanish/English condition. All other variables were the same. Because the primary effects of word type and trials have already been reported, these analyses will focus only on outcomes pertaining to language. In the first analysis, there was a marginal effect of language [$F(1,18) = 3.20$, $MS_e = 0.33$, $p < .10$] and interactions of language and word type [$F(1,18) = 61.34$, $MS_e = 2.33$, $p < .001$], language and trials [$F(4,72) = 11.17$, $MS_e = 0.30$, $p < .001$], and language, word type, and trials [$F(4,72) = 4.43$, $MS_e = 0.14$, $p < .01$]. For our purpose, the most important finding is the language $\times$ word type interaction. As shown by the grand means in Table 2, the accurate recognition of list words (.81) exceeded the false recognition of critical words (.60) in the English/English condition, whereas false recognition of critical words (.22) exceeded the false recognition of list words (.12) in the English/Spanish condition. Moreover, there was more false recognition of critical words in the same language (.60) than in a different language (.22) [$F(1,18) = 9.58$, $MS_e = 0.05$, $p < .01$].

Essentially the same results were obtained in the second analysis, which compared the Spanish/Spanish and Spanish/English conditions. There were reliable effects of language [$F(1,18) = 42.87$, $MS_e = 2.94$, $p < .001$], language and word type [$F(1,18) = 98.99$, $MS_e = 4.89$, $p < .001$], language and trials [$F(4,72) = 26.51$, $MS_e = 0.40$, $p < .001$], and language, word type, and trials [$F(4,72) = 6.54$, $MS_e = 0.15$, $p < .001$]. Again, the most important finding is the language $\times$ word type interaction. The grand means in Table 2 show that the accurate recognition of list words (.82) exceeded the false recognition of critical words (.43) in the Spanish/Spanish condition, whereas false recognition of critical words (.15) exceeded the false recognition of list words (.10) in the Spanish/English condition. Once again, there was more false recognition of critical words in the same language (.43) than in a different language (.15) [$F(1,18) = 4.78$, $MS_e = 0.03$, $p < .05$].

DISCUSSION

This study demonstrated that false memories can cross language boundaries in the DRM procedure. Bilingual participants falsely recognized critical words that were converging associates of previously heard lists when those

words were tested in the same language used during study or in a different language. Yet language was still important in this task, since the representations that gave rise to these false memories were not language independent. False recognition was greater for same-language than for different-language critical words, even for the first trial. Over trials, false memories for all critical words decreased, and when the languages were different between study and test, the false recognition of critical words approximated baseline after five trials. These results indicate that our bilingual participants employed both lexical- and conceptual-level representations to differing degrees over trials. During the initial trials, false-recognition errors were high because of the use of conceptual-level representations. At the same time, the availability of some lexical-level representations from exposure to the studied words permitted participants to make fewer different-language than same-language false recognitions. During later trials, false-recognition errors decreased substantially for both same-language and different-language critical words as a result of the increased availability of lexical-level representations and the decreased reliance on conceptual-level representations.

When Do False Memories Cross Language Boundaries?

These results support those of other bilingual studies of false memory. First, this research extends the findings of the previous single-trial study by Kawasaki-Miyaji et al. (2003) by specifying the conditions under which false memory does and does not cross language boundaries in the DRM procedure. Second, our findings are relevant to Shaw et al.'s (1997) misinformation effect study, in which memory errors were observed when misinformation was presented in the same language used in the original event or in a different language. In their study, participants perceived the meaning of the information to be noteworthy, not the specific language that provided that meaning. These findings, together with the present results, suggest the following rule: False memories can cross language boundaries in bilinguals when the language format is not deemed relevant to the task (Shaw et al., 1997) or, if the language is important, when this lexical information is not highly accessible (Kawasaki-Miyaji et al., 2003; the early trials of this study). Conversely, false memories should have difficulty crossing language boundaries when the specific language is important and lexical information is accessible (the later trials of this study).

The present results also add to the language-dependent findings of Marian and Neisser (2000), who reported greater recall of autobiographical memories in bilingual participants when the language used for recall matched the one used at the time of the autobiographical events. Our design did not permit testing for a language-dependent true-memory effect, but we did find a language-dependent false-memory effect in which more false recognition was observed when the study and test languages matched than when they differed. According to Tulving and Thomson's (1973) encoding specificity principle, retrieval is enhanced when the study and test conditions match. Whereas Mar-

ian and Neisser showed how linguistic encoding specificity was important for bilingual autobiographical recall, we have demonstrated that this concept has application to bilingual false memory: False memory is enhanced when the linguistic contexts for the study and test conditions match.

Theoretical Implications of Cross-Language False Memory

The fact that false memories crossed language boundaries, at least in the initial trials, is inconsistent with the view that information is only represented in language-specific lexicons. This view would predict false recognition of critical words only when the language was the same between study and test, not when the language was different. Instead, our results are consistent with models of bilingual representation that assume separate lexical representations for a word in each language, but common conceptual-level representations for both languages (see French & Jacquet, 2004; Kroll & de Groot, 1997; Smith, 1997). These results do not address the issue of whether our bilingual participants could directly access the conceptual-level representations from either language at test (English or Spanish) or whether access to these conceptual representations was first mediated by translation from the less proficient language (Spanish) through the more proficient language (English). In addition, these findings do not explicitly reject the view that participants merely translated word forms between separate language systems. But, as Kroll and de Groot have argued, a shift from reliance on the primary language to direct conceptual processing in the second language occurs with increasing proficiency in the second language. Given that our participants were bilingual for much of their lives, it may be that they were able to access conceptual-level representations more or less directly from either language at test.

The present findings are also in general agreement with fuzzy-trace theory (see, e.g., Brainerd & Reyna, 2002). This approach places a heavy emphasis on a memory-editing process, called *recollection rejection*, that works to diminish false recollection over trials as participants gradually reduce their reliance on conceptual-level representations (gist traces) and increase their use of lexical-level representations and other surface information (verbatim traces). In the present experiment, the specific language of a study list provided the participants with additional surface information that could aid editing by functioning as a discriminative cue to diminish false recognition. Although a test word might overlap in meaning with a studied word, if that word was not previously heard in the test language, that word should not be recognized. Gradually, participants applied this rule and reduced their false recollections over trials. Clearly, the task of bilingual participants in a mixed-language presentation of DRM lists is more complex than that of monolingual participants with single-language lists. Not only must bilingual participants distinguish between list words and nonstudied words (in-

cluding critical words), they must also distinguish between languages and remember which language was used for each list at study. Based on the present results, these distinctions are acquired only after multiple exposures to the study lists.

Yet even with multiple study-test trials, clear language differences remained. Our bilingual participants showed nearly equivalent recognition accuracy for list words over each of the five trials in the same-language condition, but they demonstrated more false recognition of critical words in the English/English condition than in the Spanish/Spanish condition (see Table 2). We suggested previously that a false-recognition difference could reflect a subtle proficiency difference in our bilingual participants. Even when participants are highly proficient in two languages, as indicated by equivalent accurate recognition for English and Spanish, performance asymmetries can be observed on other measures (see Kroll & de Groot, 1997). Intriguingly, false recognition may be a more sensitive measure of language proficiency than is accurate recognition.

Alternatively, the word lists used in this study were English lists and Spanish translations of those lists, based on the Stadler et al. (1999) norms from English speakers. False-memory norms, however, may be different for Spanish than for English DRM lists. In fact, previous research has shown that the degree to which concepts in two languages overlap can depend on their level of concreteness (see, e.g., de Groot, 1992; Kroll & de Groot, 1997), so a one-to-one conceptual correspondence between languages is not always found. In the same vein, we previously outlined the reasons why it was not possible to provide a one-to-one translation between languages for our lists. To the extent that this issue influenced the construction of the present lists, more false memory would be expected for the English lists than for the Spanish translations of those lists, which was the pattern of results we observed.

Finally, we are mindful that our results were likely influenced by our choice of participants. We employed only proficient bilingual participants, defined as having acquired both languages early in life. Given the limited transfer of false memories across language boundaries with these participants, it may be difficult to demonstrate this transfer with bilinguals who are not proficient. This suggestion follows from Brainerd, Reyna, and Forrest (2002), who reported near-floor levels of false memory in kindergarten children with monolingual DRM lists. False memory for critical words requires a certain level of language proficiency for participants to "get the gist" of the DRM lists (Brainerd et al., 2002). We argue that this finding applies also to bilingual participants. At the same time, we suspect that the ability to demonstrate cross-language false memories should not be limited to spoken languages, such as English and Spanish. It might also be observed, for example, between English and American Sign Language. The important point, we contend, is that cross-language false memory is closely linked with the use of conceptual-level language representations.

REFERENCES

- BRAINERD, C. J., & REYNA, V. F. (2002). Recollection rejection: How children edit their false memories. *Developmental Psychology*, **38**, 156-172.
- BRAINERD, C. J., REYNA, V. F., & FORREST, T. J. (2002). Are young children susceptible to the false-memory illusion? *Child Development*, **73**, 1363-1377.
- CHEN, H.-C., & NG, M.-L. (1989). Semantic facilitation and translation priming effects in Chinese-English bilinguals. *Memory & Cognition*, **17**, 454-462.
- DEESE, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, **58**, 17-22.
- DE GROOT, A. M. [B.] (1992). Determinants of word translation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **18**, 1001-1018.
- DURGUNOĞLU, A. Y., & ROEDIGER, H. L. (1987). Test differences in accessing bilingual memory. *Journal of Memory & Language*, **26**, 377-391.
- FOX, E. (1996). Cross-language priming from ignored words: Evidence for a common representational system in bilinguals. *Journal of Memory & Language*, **35**, 353-370.
- FRENCH, R. M., & JACQUET, M. (2004). Understanding bilingual memory: Models and data. *Trends in Cognitive Sciences*, **8**, 87-93.
- GARCÍA-BAJOS, E., & MIGUELES, M. (1997). Falsas memorias en el recuerdo y reconocimiento de palabras [False memory in the recall and recognition of words]. *Estudios de Psicología*, **58**, 3-14.
- GERARD, L. D., & SCARBOROUGH, D. L. (1989). Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 305-315.
- GOLLAN, T. H., & KROLL, J. F. (2001). Bilingual lexical access. In B. Rapp (Ed.), *The handbook of cognitive neuropsychology: What deficits reveal about the human mind* (pp. 321-345). Philadelphia: Psychology Press.
- KAWASAKI-MIYAJI, Y., INOUE, T., & YAMA, H. (2003). Cross-linguistic false recognition: How do Japanese-dominant bilinguals process two languages, Japanese and English? *Psychologia*, **46**, 255-267.
- KEATLEY, C. W., SPINKS, J. A., & DE GELDER, B. (1994). Asymmetrical cross-language priming effects. *Memory & Cognition*, **22**, 70-84.
- KENSINGER, E. A., & SCHACTER, D. L. (1999). When true memories suppress false memories: Effects of ageing. *Cognitive Neuropsychology*, **16**, 399-415.
- 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. de Groot and J. Kroll (Eds.), *Tutorials in bilingualism* (pp. 169-199). Mahwah, NJ: Erlbaum.
- LOFTUS, E. F., DONDERS, K., HOFFMAN, H. G., & SCHOOLER, J. W. (1989). Creating new memories that are quickly accessed and confidently held. *Memory & Cognition*, **17**, 607-616.
- MACLEOD, C. M. (1976). Bilingual episodic memory: Acquisition and forgetting. *Journal of Verbal Learning & Verbal Behavior*, **15**, 347-364.
- MARIAN, V., & NEISSER, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General*, **129**, 361-368.
- PÉREZ-MATA, M. N., READ, J. D., & DIGES, M. (2002). Effects of divided attention and word concreteness on correct recall and false memory reports. *Memory*, **10**, 161-177.
- ROEDIGER, H. L., III, & GALLO, D. A. (2002). Processes affecting accuracy and distortion in memory: An overview. In M. L. Eisen (Ed.), *Memory and suggestibility in the forensic interview* (pp. 3-28). Mahwah, NJ: Erlbaum.
- ROEDIGER, H. L., III, & McDERMOTT, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 803-814.
- SEAMON, J. G., BERKO, J. R., SAHLIN, B. H., YU, Y., COLKER, J. M., & GOTTFRIED, D. H. (in press). Can false memories spontaneously recover? *Memory*.
- SEAMON, J. G., GOODKIND, M. S., DUMEY, A. D., DICK, E., AUFSEESER, M. S., STRICKLAND, S. E., ET AL. (2003). "If I didn't write it, why would I remember it?" Effects of encoding, attention, and practice on accurate and false memory. *Memory & Cognition*, **31**, 445-457.
- SEAMON, J. G., LEE, I. A., TONER, S. K., WHEELER, R. H., GOODKIND, M. S., & BIRCH, A. D. (2002). Thinking of critical words during study is unnecessary for false memory in the Deese, Roediger, and McDermott procedure. *Psychological Science*, **13**, 526-531.
- SEAMON, J. G., LUO, C. R., KOPECKY, J. J., PRICE, C. A., ROTHSCHILD, L., FUNG, N. S., & SCHWARTZ, M. A. (2002). Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Memory & Cognition*, **30**, 1054-1064.
- SHAW, J. S., III, GARCIA, L. A., & ROBLES, B. E. (1997). Cross-language postevent misinformation effects in Spanish-English bilingual witnesses. *Journal of Applied Psychology*, **82**, 889-899.
- SMITH, M. C. (1997). How do bilinguals access lexical information? In A. de Groot and J. Kroll (Eds.), *Tutorials in bilingualism* (pp. 145-168). Mahwah, NJ: Erlbaum.
- SNODGRASS, J. G., & CORWIN, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, **117**, 34-50.
- STADLER, M. A., ROEDIGER, H. L., III, & McDERMOTT, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition*, **27**, 494-500.
- STEIN, L. M., & PERGHER, G. K. (2001). Criando falsas memorias em adultos por meio de palavras associadas [Creating false memories in adults using associated word lists]. *Psicologia: Reflexão e Crítica*, **14**, 353-366.
- TULVING, E., & THOMSON, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, **80**, 352-373.
- WATKINS, M. J., & PEYNIRCIOĞLU, Z. F. (1983). The bilingual lexicon: Language specific units in an integrated network. *Journal of Verbal Learning & Verbal Behavior*, **23**, 519-539.
- ZEELLENBERG, R., & PECHER, D. (2003). Evidence for long-term cross-language repetition priming in conceptual implicit memory tasks. *Journal of Memory & Language*, **49**, 80-94.

APPENDIX
Critical Words With List Words 1–10 for List Sets A and B

List Set A

WINDOW: door, glass, pane, shade, ledge, sill, house, open, curtain, frame
 ANGER: mad, fear, hate, rage, temper, fury, ire, wrath, happy, fight
 SWEET: sour, candy, sugar, bitter, good, taste, tooth, nice, honey, soda
 MOUNTAIN: hill, valley, climb, summit, top, molehill, peak, plain, glacier, goat
 SLOW: fast, lethargic, stop, listless, snail, cautious, delay, traffic, turtle, hesitant
 MAN: woman, husband, uncle, lady, mouse, male, father, strong, friend, beard
 OLOR (SMELL): nariz, respire, olor, aroma, oiga, ver, orificio nasal, gallo, hedor, fragancia
 DORMIR (SLEEP): cama, descanso, despierto, cansado, sueno, estela, friza, letargo, ronquido, siesta
 BASURA (TRASH): residuo, lata, desecho, alcantarillado, chatarra, desperdicios, barrer, recorte, pila, vertedero
 LADRON (THIEF): robe, dinero, policia, malo, robar, carcel, armas de fuego, bribon, crimen, banco
 MUSICA (MUSIC): nota, sonido, piano, canta, radio, grupo, melodía, trompeta, concierto, instrumento
 PAN (BREAD): mantequilla, comida, comer, bocadillo, centeno, mermelada, leche, harina, pasta, corteza

List Set B

COLD: hot, snow, winter, warm, ice, wet, frigid, chilly, heat, weather
 SOFT: hard, light, pillow, plush, loud, cotton, touch, fluffy, feather, kitten
 CHAIR: table, sit, legs, seat, couch, desk, wood, cushion, swivel, stool
 DOCTOR: nurse, sick, lawyer, medicine, health, hospital, dentist, physician, office, stethoscope
 NEEDLE: thread, pin, eye, sewing, sharp, point, pricked, thimble, thorn, hurt
 CITY: town, crowded, state, capital, streets, subway, country, New York, village, metropolis
 DURO (ROUGH): suave, abollado, camino, aspero, papel de lija, dentado, grueso, desigual, escabroso, arena
 TAZA (CUP): jarro, platillo, té, tapa, mango, café, paja, copa, sopa, bebida
 HUMO (SMOKE): cigarillo, llama, polucion, ceniza, cigarro, chimenea, fuego, tabaco, heder, pipa
 ALTO (HIGH): bajo, nubes, arriba, torre, salto, sobre, edificio, precipicio, cielo, encima
 RIO (RIVER): agua, arroyo, lago, Misisipí, barco, marea, nadar, flujo, barcaza, cala
 GOMA (RUBBER): elastico, rebote, llanta, bola, borrador, espuma, planta, latex, cola, flexible

(Manuscript received August 22, 2003;
 revision accepted for publication January 5, 2005.)