APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **20**: 1025–1047 (2006)
Published online 21 June 2006 in Wiley InterScience
(www.interscience.wiley.com). DOI: 10.1002/acp.1242

Language-Dependent Memory in Bilingual Learning

VIORICA MARIAN[1]* and CAITLIN M. FAUSEY[2]

[1]*Northwestern University, USA*
[2]*Stanford University, USA*

SUMMARY

Spanish-English bilinguals were taught academic-type information about History, Biology, Chemistry and Mythology in their two languages. Upon testing, it was found that memory was more accurate and retrieval was faster when the language of retrieval and the language of encoding matched than when they did not match. For accuracy, the pattern of results was influenced by bilinguals' language proficiency, so that only balanced bilinguals whose high proficiency levels were similar in both languages showed language-dependent recall. For reaction time, bilinguals were faster to retrieve information when the languages of retrieval and encoding matched than when they mismatched, but only for material encoded in Spanish. The influence of encoding and retrieval languages on error patterns was also examined. Together, the study's findings suggest that bilingual learning may be subject to language dependency and that experience with a language may increase the strength of linguistic cues in producing language-dependent memory. The results are consistent with previous findings of language-dependent memory in autobiographical narratives, carry applied implications for bilingual education, and are discussed within the theoretical framework of the relationship between language and memory. Copyright © 2006 John Wiley & Sons, Ltd.

The relationship between language and memory is embedded in theoretical discussions about the extent to which thought is linguistic (e.g. Chomsky, 1980; Hespos & Spelke, 2004; Pinker, 1994; Whorf, 1956) and the degree to which cognitive processes vary according to the language one speaks (e.g. Boroditsky, Ham, & Ramscar, 2002; Gennari, Sloman, Malt, & Fitch, 2002; Pederson et al., 1998). The influences of language on conceptual representations have also been considered in bilingualism research (see Francis, 1999b; Heredia & McLaughlin, 1992; Kroll & de Groot, 1997 for reviews). The present work contributes to understanding the relationship between language and memory by studying bilingual learning.

We begin with the assumption that at least some aspects of mental processes are connected to language, a notion supported by evidence that linguistic expression influences memory (e.g. Billman & Krych, 1998; Billman, Swilley, & Krych, 2000; Feist & Gentner, in press; Loftus & Palmer, 1974). In monolingual studies, using path and manner verbs

*Correspondence to: Dr. V. Marian, Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL 60208-3570, USA. E-mail: v-marian@northwestern.edu

Contract/grant sponsor: NICHD; contract/grant number: 1R03HD046952-01A1; contract/grant number: NSF; BCS-0418495.

prior to (Billman et al., 2000) or during (Billman & Krych, 1998) watching a videotaped event was found to influence subsequent recognition of old and new events. Similarly, viewing ambiguous spatial representations paired with spatial prepositions (e.g. 'on') resulted in a false memory bias towards typical portrayals of the spatial representation (Feist & Gentner, in press). Even manipulating language at retrieval only was found to influence memory for previously encoded information (e.g. Loftus & Palmer, 1974). For instance, using the word 'smashed' in questions about the speed of a car in a previously-viewed videotape resulted in higher speed estimates and more reported false memories of broken glass than using the word 'hit.'

While the general notion that language use during encoding and retrieval impacts memory is widely accepted, the exact ways in which languages at encoding and retrieval influence memory are less understood. Previous research suggests that memory for information learned in a particular environment is improved when the retrieval context is similar to the original encoding context (see Davies & Thomson, 1988; Tulving & Thomson, 1973, for a review). This encoding-specificity principle was later extended to linguistic context when it was found that bilinguals' autobiographical memory was facilitated by using the same language at retrieval as had been used during the original encoding (Marian & Neisser, 2000). Moreover, bilinguals were found to exhibit more intense emotion when the language at retrieval matched the language at encoding than when the two did not match (Marian & Kaushanskaya, 2004), and to spontaneously switch languages more often when describing events that happened in the other language (e.g. Marian & Kaushanskaya, 2005; Otheguy & García, 1993).

In addition, several experiments that focused on bilingual memory for word lists (e.g. Durgunoglu & Roediger, 1987; Kintsch, 1970; list. –>Watkins & Peynircioglu, 1983) found greater facilitation of recall and recognition when no language switch took place between study and test than when the languages did switch. In studies that presented mixed-language word lists to bilinguals (e.g. Dalrymple-Alford & Aamiry, 1969; Lambert, Havelka, & Crosby, 1958; Nott & Lambert, 1968), recall output was organized by both language and semantic category. Analyses by Francis, (1999b) suggested that the degree to which words were recalled in same-language clusters was significantly greater than the language clustering that had been present in the studied lists. Because strategies such as covert translation can present problems in studies of cross-language recall and recognition, effects of language-switching have also been studied using tests of implicit memory. Several studies, for instance, have found that repetition priming in paired-associate learning was greater when the same language was maintained across sessions (e.g. Durgunoglu & Roediger, 1987; Heredia & McLaughlin, 1992; Macleod, 1976; Peynircioglu & Durgunoglu, 1993; Smith, 1991). However, results of bilingual word-list and implicit memory studies, as well as of picture naming (e.g. Ervin, 1961) and problem-solving (e.g. Bernardo, 1998; Francis, 1999a) studies, have not always been consistent. Variability in methodology (Marian, 2006) may have resulted in different processing mechanisms being tapped into (see Heredia & McLaughlin, 1992; Kroll & de Groot, 1997 for reviews) and are likely to be responsible for these differences. The exact circumstances under which there is a processing advantage for same-language study-test pairs remain unclear, underscoring a need to pinpoint the specific types of learning and memory that are susceptible to language dependency and the mechanisms driving such effects.

Given the findings of language-dependent autobiographical memory in bilinguals, of processing advantages in same-language test-retest pairs of wordlists, and of phonological facilitation in monolinguals if similar features are used at encoding and retrieval (e.g. Bradlow, Nygaard, & Pisoni, 1999; Nygaard, Burt, & Queen, 2000; Singh, Morgan, & White, 2004), it seems likely that language overlap at encoding and retrieval may improve memory access for ecologically-valid semantic material as well. One context that is largely linguistic is the educational classroom, where the influence of language is manifested at the critical endpoints of the learning continuum: information is presented linguistically by the teacher, and the student often must use language to demonstrate that s/he has learned the information. Understanding the role of language in the encoding and retrieval of information, then, is important at least for the stages of communication central to the educational endeavor. From this perspective, the relationship between language and memory in explicitly linguistic contexts is not trivial. Not only would finding language-dependent memory have applied implications for educational testing and clinical services, but it may also inform theoretical discussions about the role of linguistic context in the representation and processing of semantic knowledge.

The goals of the present paper, then, were to examine the presence of language-dependent memory patterns in learning of academic material and to propose mechanisms that may underlie language-dependent memory. Bilingual speakers of Chilean Spanish and English were taught academic-type information in either Spanish or English and memory was tested in the two languages. In studies of linguistic and cognitive processing in bilinguals, variables such as language proficiency, dominance and experience have been found to influence patterns of results (Grosjean, 1997; Marian, 2006). For example, evidence suggests that as proficiency levels change, so do the organization and processing of the two languages (e.g. Kroll & Stewart, 1994). To account for these potential influences, a detailed questionnaire was used to assess bilinguals' linguistic profiles (Marian, Blumenfeld, & Kaushanskaya, 2005).

We predicted that proficiency in the two languages would mediate the relationship between language and memory and that this influence would be driven by similarities and differences in processing the two languages. Namely, bilinguals with comparable proficiency levels in their two languages may employ similar cognitive strategies and processing resources in both their first and second languages, while bilinguals with varying degrees of proficiency in their two languages may employ different strategies and resources in their first and second languages. Consequently, the effect of proficiency on language-dependent memory patterns may manifest itself in two possible ways. On the one hand, language-dependent memory may be magnified when processing differences between languages are greater, because a mismatch in languages is also associated with a mismatch in many other variables. By that account, the more similar the proficiency levels in two languages, the *weaker* the language-dependent memory effects. On the other hand, if language itself serves as a cue, then language-dependent memory may be magnified when no other cues differ between encoding and retrieval and do not compete for salience with the language cue. By that account, the more similar the proficiency levels in two languages, the *stronger* the language-dependent memory effects. Because it is unclear what role processing differences play in language-dependent memory, the direction of the effect remains an open empirical question, one that the present experiment aimed to inform. Because it is never the case that bilingual samples are entirely homogeneous with respect to linguistic profile, our study had the potential to reveal how differences such as relative proficiency in the two languages influence language-dependent memory. By either account, differences in the ability to understand material across languages and differences in processing a higher- versus a lower-proficiency language were predicted to influence language-dependent memory phenomena. In sum, our hypotheses were as follows:

(a) Bilinguals would produce more correct answers when responding to questions asked in the same language in which the information was originally learned than in the other language;
(b) Bilinguals would respond faster when the languages of encoding and retrieval matched than when they mismatched;
(c) Bilinguals' proficiency in their two languages would influence their susceptibility to language-dependent memory effects.

METHOD

Participants

Twenty-four Chilean Spanish-English bilinguals (9 males, 15 females) participated in the study (two other participants guessed the hypothesis of the study and were excluded from analyses). Participants' mean age at the time of the experiment was 22 years ($SD = 2.93$ years). All participants lived in a Spanish-speaking country (Chile) at the time of testing. Participants were native speakers of Spanish who learned English between the ages of 0 and 12 years ($M = 4.00$ years, $SD = 2.52$ years) and were fluent in both Spanish and English. Participants' language proficiency (understanding, speaking and reading[1]), as well as age of acquisition, current use and cultural affiliation were assessed using the Language Experience And Proficiency Questionnaire (LEAP-Q, Marian et al., 2005).

Self-reported ratings of proficiency were collected on a scale from 5 (high) to 0 (none) for both Spanish and English. Across all participants, analyses revealed higher self-reported levels of proficiency understanding Spanish ($M = 4.92$, $SD = 0.28$) than English ($M = 4.33$, $SD = 0.56$), $t(23) = 4.37$, $p < 0.0001$, higher levels of proficiency speaking Spanish ($M = 4.83$, $SD = 0.38$) than English ($M = 4.04$, $SD = 0.91$), $t(23) = 3.65$, $p = 0.001$, and higher levels of proficiency reading Spanish ($M = 4.83$, $SD = 0.48$) than English ($M = 4.21$, $SD = 0.78$), $t(23) = 3.32$, $p = 0.003$. In addition, proficiency speaking was independently coded by an English-Spanish bilingual using audio-recordings and confirmed higher levels of proficiency speaking Spanish ($M = 5.00$, $SD = 0.00$) than English ($M = 3.63$, $SD = 1.06$), $t(23) = 6.38$, $p < 0.00001$.

Early analyses of LEAP-Q data revealed that the bilingual sample tested was not homogeneous and varied in levels of relative proficiency understanding Spanish and English. Because the ability to understand information is critical in the encoding process, and because similarities and differences in proficiency levels across the two languages may influence language-dependent memory effects, analyses were performed in which participants were grouped into a *balanced-bilinguals* group and an *unbalanced-bilinguals* group, based on self-reported proficiency understanding Spanish and English. Proficiency understanding was used for two reasons—(1) because the task relied most heavily on the ability to understand the stories and questions (as opposed to proficiency speaking, which did not play any role during encoding, or proficiency reading and writing, which were not targeted in this experiment); and (2) because proficiency

[1]The Language Experience and Proficiency Questionnaire does not assess self-reported proficiency writing, due to previous findings in the course of questionnaire development that self-reported proficiency writing correlates highly with self-reported proficiency reading and that proficiency reading is a better predictor of overall language proficiency as revealed by factor analyses and behavioral measures (Marian et al., 2005).

understanding is a particularly reliable and valid self-reported measure, as evident from finding that correlations between self-reported proficiency understanding and performance on language subtests of the Woodcock-Johnson were higher than those for any of the other self-reported proficiency measures assessed with the LEAP-Q (Marian et al., 2005).

Participants who reported higher proficiency in Spanish than in English were grouped as *unbalanced-bilinguals* ($N = 14$), and participants who reported equal levels of proficiency in the two languages were grouped as *balanced-bilinguals* ($N = 10$). For the *balanced-bilinguals* group, participants' mean proficiency understanding was 4.8 for Spanish ($SD = 0.42$) and 4.9 for English ($SD = 0.32$), a difference that was not significant. For the *unbalanced-bilinguals* group, participants' mean proficiency understanding was 5.0 for Spanish ($SD = 0.00$) and 3.93 for English ($SD = 0.27$), $t(13) = 15.00$, $p < 0.0001$.

Table 1 provides the means and standard deviations for all relevant language background measures collected with the Language Experience and Proficiency Questionnaire. It also indicates statistically significant differences between Spanish and English within each bilingual group, as well as statistically significant differences between balanced and unbalanced bilinguals. For each reported measure, a 2×2 Analysis of Variance, with group (balanced, unbalanced) as a between-subjects variable and language (Spanish, English) as a within-subjects variable, was conducted. For each proficiency measure, the interaction between group and language was significant. Follow-up comparisons were conducted and analyses that yielded significant differences are marked with asterisks in Table 1. In all cases, the difference between Spanish and English proficiency measures was smaller for balanced bilinguals than for unbalanced bilinguals. Specifically, *t*-tests for measures of proficiency revealed that while unbalanced bilinguals reported significantly different levels of proficiency understanding, speaking, and reading

Table 1. Language history for balanced and unbalanced bilinguals, based on self-reported data. (The proficiency scale ranged from 0 (none) to 5 (high). Ages and length of time are reported in years.)

	Balanced bilinguals		Unbalanced bilinguals	
Measure	Spanish (*SD*)	English (*SD*)	Spanish (*SD*)	English (*SD*)
Mean age when tested	21.5 (2.59)		22.4 (3.2)	
Proficiency understanding	4.8 (0.42)	4.9 (0.32)†	5.0 (0.00)**	3.9 (0.27)
Proficiency speaking	4.6 (0.52)	4.5 (0.71)†	5.0 (0.00)**,†	3.7 (0.91)
Proficiency reading	4.7 (0.67)	4.6 (0.70)†	4.9 (0.27)**	3.9 (0.73)
Age when began learning	0.95 (1.01)*	4.2 (3.29)	0.96 (1.25)**	3.9 (1.93)
Age when fluent	4.4 (1.89)	8.6 (5.75)	5.0 (2.87)**	9.6 (3.83)
Years in country where language spoken	19.1 (4.22)**	2.4 (2.90)†	22.1 (3.33)**	0.6 (0.83)
Years in family where language spoken	19.0 (7.41)*	6.8 (10.48)	22.4 (3.14)**	8.2 (11.44)
Years in school/work where language spoken	14.1 (6.48)	8.8 (5.41)	18.8 (6.13)**	7.8 (6.90)
Current exposure (per cent)	60.0 (12.91)*	34.9 (10.06)†	78.6 (11.99)**,†	19.6 (11.27)

$^{*}p < 0.05$, Spanish versus English paired-samples *t*-test.
$^{**}p < 0.01$, Spanish versus English paired-samples *t*-test.
$^{\dagger}p < 0.05$, Balanced versus Unbalanced Bilinguals independent-samples *t*-test. Symbol location indicates the group with the higher mean value.

across the two languages (at $p < 0.001$), balanced bilinguals' ratings on these measures did not differ significantly ($p > 0.05$).

Materials

The materials for this study were designed so as to be ecologically valid and representative of the type of academic material a student might encounter in an actual classroom. Four short stories about mythology, history, biology, and chemistry and four corresponding sets of questions were prepared. The English version of the stories and questions is included in the Appendix, with the Spanish version available upon request. The four stories were constructed so as to contain fictitious but reasonable information in order to make the content meaningful, while at the same time preventing previous knowledge of the material. The mythology story described a myth associated with celebrating the beginning of winter held by a fictitious group of people. The history story described the causes, course, and consequences of a war between two fictitious nations. The biology story described the flora of a fictitious island. The chemistry story described the accidental discovery and properties of a fictitious chemical element. All four stories were balanced in length within and across languages. For each story, 10 questions were prepared,[2] constructed in such a way so as not to prime for answers to other questions.

The English version of all stories and questions was translated into Spanish by an English-Spanish bilingual and the Spanish version of each story was then verified by two Spanish-English bilinguals who were native speakers of Chilean Spanish. To control for variability in the experimenter's presentation of stimuli and questions (e.g. intelligibility, speed, etc.), after the final version of the stimuli was developed, all materials were tape-recorded in a soundproof booth. Care was used in selecting the speaker to make the recording. Several bilingual speakers of Chilean Spanish and English were contacted to identify highly proficient speakers of both languages, and recordings of speech samples were made for three Spanish-English bilinguals from Chile. Eleven independent judges rated the three speech samples for intelligibility and absence of foreign accent in English and the speaker who received the highest ratings was selected.

Design

The study followed a 2×2 repeated-measures factorial design, with language of encoding (Spanish, English) and language of retrieval (Spanish, English) as within-subjects independent variables. Participants listened to four stories, two in Spanish and two in English. The Chemistry and Mythology stories were always presented in the same language, as were the Biology and History stories, so that one science-oriented story and one humanities-oriented story were presented in each language.

Participants listened to two blocks of stories (one block of two stories in Spanish, one block of two stories in English), and two blocks of questions (one block of 20 questions in Spanish, one block of 20 questions in English). Each block of questions contained five questions per story, thus, five questions matched the language in which the story was read and five questions did not match the story language. Questions within each block were

[2]Two more questions, targeting lexical information (e.g. names) were also included as a pilot component for future research on memory for different types of information. The low number of items ($n = 2$) did not permit statistical analyses of those data.

grouped by story, so that the five questions about one story were presented together. Between each block, participants completed a short, timed puzzle selected from a puzzle book (Barber, 2001).[3] Thus, the order of the study was as follows:

(1) Block of two stories in one language (Counterbalanced)
(2) Puzzle completion distracter task
(3) Block of two stories in the other language (Counterbalanced)
(4) Puzzle completion distracter task
(5) Block of 20 questions in one language (Counterbalanced)
(6) Puzzle completion distracter task
(7) Block of 20 questions in the other language (Counterbalanced)

Language of instruction and language of testing were counterbalanced across participants, so that half of all participants heard stories in Spanish first and the other half heard stories in English first. Of the 12 participants who heard the Spanish stories first, six were presented with the Spanish questions before the English questions and six were presented with the English questions before the Spanish questions. Similarly, of the 12 participants who heard the English stories first, six were presented with the Spanish questions before the English questions and six were presented with the English questions before the Spanish questions. Thus, the language order was as follows:

(1) Six participants heard English stories first, followed by Spanish stories, and were tested with English questions first, followed by Spanish questions;
(2) Six participants heard English stories first, followed by Spanish stories, and were tested with Spanish questions first, followed by English questions;
(3) Six participants heard Spanish stories first, followed by English stories, and were tested with Spanish questions first, followed by English questions;
(4) Six participants heard Spanish stories first, followed by English stories, and were tested with English questions first, followed by Spanish questions.

Finally, the language in which each particular story was presented was also counterbalanced. Within each of the four language order groups described, half of the participants were presented with the Chemistry and Mythology stories in Spanish and the Biology and History stories in English, and the other half of participants were presented with the Chemistry and Mythology stories in English and the Biology and History stories in Spanish. While this counterbalancing was planned for each language order group, the testing protocol was accidentally switched for one participant in each of the language order groups that began with English stories. To ensure that this counterbalancing error did not impact overall results, we performed *post hoc* analyses for each story to examine potential per-story differences in response patterns due to the language of story presentation, language of question presentation, language matching and language mismatching. For each story, performance was compared for participants who heard that story in Spanish versus in English, retrieved it in Spanish versus in English, and were presented with the match Spanish-Spanish condition versus the match English-English condition, and with the mismatch Spanish-English condition versus the mismatch English-Spanish condition.

[3]The more traditional non-linguistic distracter task of solving numerical math problems typically used with monolinguals was not used because it may bias the bilinguals towards the language in which mathematical knowledge was originally acquired (e.g. Spelke & Tsivkin, 2001). Timing the puzzle completion using a stopwatch was intended to increase the effectiveness of the puzzle as a distracter task.

These *post hoc* analyses showed no significant differences (all p values were greater than 0.2). Because an equal number of participants answered questions in matched (languages of encoding and retrieval were the same) and in mismatched (languages of encoding and retrieval were different) conditions for each story, and because specific language contrasts per story did not differ, the validity of the results is unlikely to have been impacted by the minor counterbalancing error.

Prior to testing, participants were instructed to answer all questions in the same language in which the questions were asked and not to switch languages at any point. Participants' responses were recorded and their language of response was noted. Participants consistently provided their answers in the target language and without switching to the other language. Across all participants and questions, only three instances of code-switching (using the non-target language when answering questions in a target language) were recorded.[4] All three code-switches took place in the mismatch condition, when material encoded in Spanish was being retrieved in English.

Procedure

After completing consent forms in both languages, participants listened to instructions in the language in which the first two stories were presented. The instructions were provided in a face-to-face conversation between an English–Spanish bilingual experimenter and the participant. Participants were instructed to remember information from the stories, and were told that they would later answer questions about the stories. No mention was made about the language of the questions, about the relationship between story language and question language, or about the fact that, regardless of story language, questions would be in both English and Spanish. Instead, participants were told that the purpose of the experiment was to examine bilingual performance on non-verbal spatial tasks after learning information in one language versus another. This foil was employed in order to discourage participants from focusing on the relationship between the language of the stories and the language of the questions, as well as to increase the effectiveness of the puzzles as a distracter task. Later analyses showed no differences in puzzle-completion time depending upon language used immediately prior to completing a puzzle. Bilinguals took about 135 seconds to complete the puzzle after hearing Spanish and 129 seconds after hearing English, $t(25) = 0.45$, $p = 0.66$.

All stories and questions were presented through headphones on a Dell Inspiron 5000 laptop, using Windows Media Player. Participants' responses to questions were recorded using a Sony microphone and minidisc recorder. After answering all questions, participants were queried about the purpose of the study and asked to complete the Language Experience and Proficiency Questionnaire.

Coding and analyses

Dependent measures

Response accuracy, latency, and error rates were compared across each of the four encoding-language by retrieval-language conditions (Spanish encoding-Spanish retrieval, Spanish encoding-English retrieval, English encoding-Spanish retrieval and

[4]Of these, two were yielded by the same participant, who used the cognate word 'vomitos' instead of 'vomiting' and used the Spanish exclamation 'chuta' (the equivalent of 'shoot'), while engaging in self-talk and saying 'I don't know . . . chuta.' In the third case, a participant said 'elementos ocultos' instead of 'hidden elements' ('elementos' and 'elements' are also cognates).

English encoding-English retrieval). For each of the four conditions, the percentage of correct answers and the mean response latency of correct answers were calculated per participant. A second independent rater coded 12.5% of data for reliability purposes. For response accuracy, point-to-point agreement between the two coders was 93%. For response latency, point-to-point agreement between the two coders was 98% (at 1 second), 96% (at 0.5 seconds) and 84% (at 0.1 seconds). In addition, *post hoc* coding of all errors into errors of omission, errors of commission and partial errors was performed for further error analyses across conditions.

Response accuracy

Responses were coded as correct if the participant provided information that answered the question and was consistent with the story. For each participant, the raw number of correct answers in each of the four encoding language-retrieval language conditions was converted to a percentage correct score (raw number correct answers/total questions). All errors were content-related. Except for the three instances discussed earlier, no language errors in which participants switched to the language other than the question language took place.

Response latency

Response latency was defined as the time between the end of a spoken question and the onset of the content words of a participant's answer and was coded using Praat speech analysis software. For each participant, the mean response latency for correct answers in each of the four encoding language-retrieval language conditions was calculated. Fourteen responses (1.2% of the data) were excluded from response latency analyses due to technical errors with playing stimuli or recording answers.

Error patterns

Questions that were not answered correctly were included in an error analysis. Errors were coded as one of three kinds: (1) Errors of omission, in which participants responded 'I don't know' or 'I don't remember', (2) Errors of commission, in which participants provided an incorrect answer and (3) Partial errors, in which some correct information was provided, but participants either left out the crucial part of the answer or also included false information. For each participant, the mean number of errors of omission, the mean number of errors of commission, and the mean number of partial errors relative to the total number of answers was computed for each of the four encoding language/retrieval language conditions.

RESULTS

Response accuracy

The percentage of correct answers per condition was calculated for each participant. Data were analysed with a $2 \times 2 \times 2$ Analysis of Variance, with language of encoding (Spanish, English) and language of retrieval (Spanish, English) as within-subjects variables, and proficiency group (balanced, unbalanced) as a between-subjects variable.[5] Results revealed a significant interaction among language of encoding, language of retrieval, and proficiency, $F(1, 22) = 8.11$, $MSE = 0.027$, $p = 0.01$. No other significant main effects or interactions were found. Two planned-comparison ANOVAs, conducted separately for each proficiency group, revealed a significant interaction between language of encoding and language of retrieval in the balanced-bilinguals group, $F(1, 9) = 11.67$, $MSE = 0.015$, $p = 0.01$, but not in the unbalanced-bilinguals group, $F(1, 13) = 1.52$, $MSE = 0.035$, $p = 0.24$. No significant main effects of encoding language or retrieval language were found in either group. Follow-up paired sample *t*-tests were conducted for balanced bilinguals and results showed a strong pattern of language-dependent memory. Balanced bilinguals showed better retrieval of material encoded in Spanish when retrieved in Spanish (mean = 73%, $SE = 6\%$) than when retrieved in English (mean = 60%, $SE = 6\%$), $t(9) = 2.44$, $p = 0.04$. Similarly, balanced bilinguals showed better retrieval for material encoded in English when retrieved in English (mean = 67%, $SE = 6\%$) than when retrieved in Spanish (mean = 53%, $SE = 8\%$), $t(9) = 2.34$, $p = 0.04$ (see Figure 1).

Response latency

The mean response latencies (in seconds) of *correct* answers per condition were calculated for each participant. In 2.8% of the data, response times exceeded 2.5 standard deviations

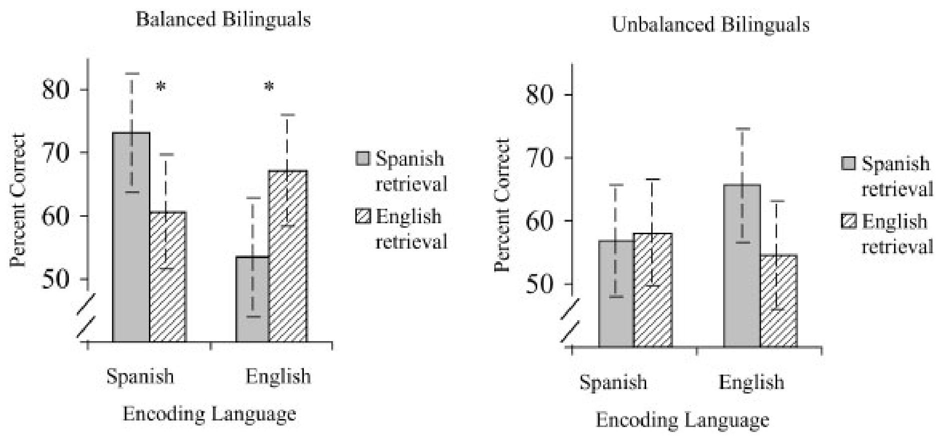

Figure 1. Response accuracy for balanced and unbalanced bilinguals. Broken lines represent a 95% confidence interval, computed following the Masson and Loftus (2003) method for repeated-measures comparisons

[5]Data were also analysed with an overall ($N = 24$) 2×2 language of encoding (Spanish, English) by language of retrieval (Spanish, English) Analysis of Covariance, where the covariate consisted of a ratio of proficiency understanding English over proficiency understanding Spanish. Results revealed a significant interaction between language of encoding and language of retrieval, $F(1, 22) = 7.70$, $MSE = 0.026$, $p = 0.01$, with higher accuracy rates when the languages of encoding and retrieval matched (mean = 62%, $SE = 4\%$) than when they mismatched (mean = 60%, $SE = 4\%$), $F(1, 22) = 7.71$, $MSE = 0.013$, $p = 0.01$ and no other significant effects. Because the covariate was not normally distributed, but rather followed a bimodal distribution, the Results section only reports comparisons in which proficiency group (balanced/unbalanced) was also included in analyses; these comparisons are more precise and provide a more accurate picture of the findings.

above the mean for that condition. Because reaction time analyses are sensitive to outliers, response times in those cases were replaced with the value of 2.5 standard deviations above the mean for that condition.

A $2 \times 2 \times 2$ Analysis of Variance with proficiency (balanced vs. unbalanced bilinguals) as a between-subjects variable and with language of encoding and language of retrieval as within-subjects variables revealed a significant interaction between languages of encoding and retrieval, $F(1, 22) = 7.61$, $MSE = 0.61$, $p = 0.01$, as well as a main effect of language of encoding $F(1, 22) = 6.92$, $MSE = 0.42$, $p = 0.02$, and of language of retrieval, $F(1, 22) = 5.60$, $MSE = 0.70$, $p = 0.03$. Neither the main effect of proficiency nor the three-way interaction among encoding language, retrieval language and proficiency were significant. Overall, participants were faster when the language of retrieval was Spanish ($M = 2.02$, $SE = 0.13$) than when it was English ($M = 2.46$, $SE = 0.21$), and when the language of encoding was English ($M = 2.05$, $SE = 0.13$) than when it was Spanish ($M = 2.43$, $SE = 0.19$). Follow-up paired-samples t-tests revealed that bilinguals showed language-dependent memory when the language of encoding was Spanish, and were faster to answer when the encoding and retrieval languages matched ($M = 1.98$, $SE = 0.13$) than when they mismatched ($M = 2.88$, $SE = 0.30$), $t(23) = 3.36$, $p = 0.003$ (see Figure 2). When the encoding language was English, response latencies did not differ significantly between the match ($M = 2.04$, $SE = 0.18$) and the mismatch conditions ($M = 2.05$, $SE = 0.14$), $t(23) = 0.07$, $p = 0.94$. Proficiency in the two languages did not influence response latencies and no differences in reaction times were found between balanced and unbalanced bilinguals.

Error analysis

Because an analysis of total errors is inversely related to accuracy analyses, only error-type analyses are reported. Across all conditions, 62% of errors were errors of omission, 22% of errors were errors of commission, and 16% were partial errors. The different error types were analysed using analyses of variance, with encoding language and retrieval language as within-subjects variables and proficiency group as a between-subjects variable. Only significant main effects and interactions are reported; absent results indicate that no significant differences were found. Means and standard errors (for each error type, in each

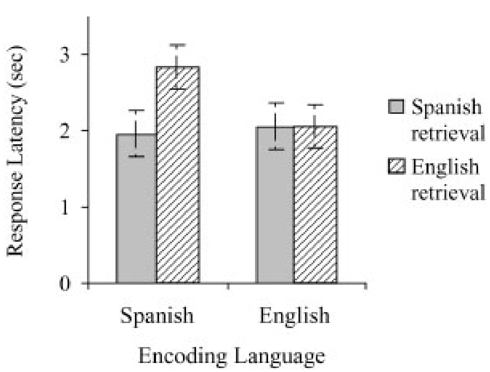

Figure 2. Response latency for all bilinguals. Broken lines represent a 95% confidence interval, computed following the Masson and Loftus (2003) method for repeated-measures comparisons

encoding language-retrieval language condition, and for each proficiency group) are shown in Table 2.

For *errors of omission*, a significant interaction among encoding language, retrieval language, and proficiency was found, $F(1, 22) = 5.90$, $MSE = 0.036$, $p = 0.02$. Follow-up 2×2 ANOVAs for each proficiency group revealed that unbalanced bilinguals made more omission errors when retrieving in English than in Spanish, $F(1, 13) = 5.35$, $MSE = 0.019$, $p = 0.04$ (see Table 2). Balanced bilinguals showed a marginal interaction between encoding and retrieval languages, $F(1, 9) = 4.75$, $MSE = 0.023$, $p = 0.057$; the higher number of omission errors was marginally significant in the Spanish-English versus Spanish-Spanish contrast, $t(9) = 2.01$, $p = 0.076$, but did not approach significance in the English-Spanish versus English-English contrast, $t(9) = 1.83$, $p = 0.10$. For *errors of commission*, the interaction between encoding language and retrieval language was significant across both groups, $F(1, 22) = 5.00$, $MSE = 0.006$, $p = 0.04$. Follow-up analyses suggested that bilinguals made more errors of commission when the encoding and retrieval languages mismatched than when they matched; the difference was marginally significant when English was the encoding language, $t(23) = 2.04$, $p = 0.05$; but did not reach significant when Spanish was the encoding language, $t(23) = 1.54$, $p = 0.14$. For *partial errors*, a marginal effect of proficiency was observed, with unbalanced bilinguals producing more partial errors than balanced bilinguals, $F(1, 22) = 4.32$, $MSE = 0.003$, $p = 0.05$.

DISCUSSION

To examine the influence of language on bilingual learning, memory performance was compared across conditions in which language at encoding and language at retrieval were varied. Spanish-English bilinguals exhibited language-dependent memory as indicated by increased accuracy, shorter response times and fewer errors when encoding and retrieval languages matched than when they mismatched. This general language-dependent memory pattern, however, was moderated by language experience variables, such, as relative proficiency in the first and second languages. The specific patterns uncovered are discussed below, along with potential mechanisms that may be driving language-dependent memory in bilinguals.

Response accuracy

Higher response accuracy rates when the languages of encoding and retrieval matched than when they mismatched were observed only in balanced bilinguals whose high levels of proficiency were similar across the two languages. Finding language-dependent memory patterns only in the balanced bilinguals is consistent with previous findings that highly proficient balanced bilinguals are more susceptible to language-dependent memory than bilinguals with varying proficiency levels when it comes to variables such as emotional intensity in bilingual autobiographical retrieval (Marian & Kaushanskaya, 2004).

Another measure of response accuracy relied on analysis of error rates. Because our study was not designed to examine different error types in bilingual memory, interpretations of error patterns are speculative and future research is needed to examine language-dependent memory patterns for various error types. Studies that are designed to encourage specific types of errors (e.g. false memory errors, incomplete answers, etc.) are necessary in order to fully understand how the match and mismatch of encoding and

Table 2. Error patterns of balanced and unbalanced bilinguals. (Values represent the mean percentage and standard errors in each of the four encoding language-retrieval language conditions)

	Balanced bilinguals				Unbalanced bilinguals			
Error type	Spanish Encoding Spanish Retrieval	Spanish Encoding English Retrieval	English Encoding Spanish Retrieval	English Encoding English Retrieval	Spanish Encoding Spanish Retrieval	Spanish Encoding English Retrieval	English Encoding Spanish Retrieval	English Encoding English Retrieval
Omission errors	13.6 (3.6)	25.2 (5.2)	30.4 (5.9)	21.0 (4.6)	26.6 (5.4)	26.8 (4.9)	18.9 (4.1)	35.8 (5.8)
Commission errors	9.0 (4.1)	10.7 (2.7)	11.0 (3.1)	7.3 (2.7)	6.4 (2.0)	11.1 (2.3)	8.7 (3.0)	3.8 (1.4)
Partial errors	4.1 (3.1)	3.7 (2.0)	5.1 (2.3)	5.0 (2.2)	10.1 (2.1)	4.4 (1.4)	6.7 (2.1)	6.1 (1.5)
Total errors	26.7	39.6	46.5	33.3	43.1	42.3	34.3	45.7

retrieval languages influence error rates. Nevertheless, a basic analysis of error patterns revealed differences across bilingual groups and error types. For errors of omission, only balanced bilinguals showed an increased error rate when the languages of encoding and retrieval did not match than when they matched. Unbalanced bilinguals produced more errors of omission when retrieving in English, their less proficient language, than when retrieving in Spanish, their more proficient language. For errors of commission, both groups produced similar patterns and exhibited language-dependent memory for material learned in English, but not for material learned in Spanish. The lowest number of commission errors was produced in the English encoding-English retrieval condition, suggesting that bilinguals may be more comfortable speculating when the learning and testing context involves their first language. Finally, neither group showed language-dependent memory for partial errors, likely due to the low overall incidence of partial errors and to the various sub-types of errors included (the low incidence rates prevented further classification of partial errors into more detailed categories for statistical analyses). Potential explanations for accuracy differences between balanced and unbalanced bilinguals are considered later in the discussion.

Response latency

While accuracy data reflect what material was learned out of the total information presented, response latency data were limited to information that was accessed *correctly* (response latencies for incorrect answers are difficult to interpret and were not considered). Reaction time analyses revealed a similar pattern for both balanced and unbalanced bilinguals, with faster response times in the matched than in the mismatched condition when the encoding language was Spanish. When the encoding language was English, the retrieval language did not seem to have as great an influence on the speed of memory access. We suggest that this similarity in response times between the English–English and the English–Spanish conditions may be the result of two different mechanisms: The English–English condition reflects the language-dependent advantage of a match in encoding and retrieval languages and the English–Spanish condition reflects the linguistic experience of bilinguals in our sample. Note that the two conditions in which the languages of encoding and retrieval matched (English–English and Spanish–Spanish) resulted in similar reaction times, while the two conditions in which the languages of encoding and retrieval did not match (Spanish–English and English–Spanish) did not. Because the range of contexts in which the bilinguals in our sample used Spanish varied more than the contexts in which they used English, we suggest that they may have had more experience retrieving information in Spanish that they had originally encoded in English than the other way around and that this experience in transferring information from one language to the other influenced the pattern of results. The response latency data suggest that experience switching between linguistic contexts may improve the speed with which bilinguals report correct information that they accessed in mismatched contexts. This hypothesis about the role of experience with mismatched language contexts is further supported by the asymmetry in response times reported by Marian and Neisser (2000). In their study of autobiographical memory, Russian–English bilinguals were slower retrieving memories encoded in English and retrieved in Russian than the other way around, a pattern that may also reflect relative experience with mismatched linguistic contexts. The bilinguals in Marian and Neisser's study had immigrated to the United States around the age of thirteen and were in their early twenties at the time of the experiment. It is likely that these bilinguals had more opportunities to express autobiographical memories that had been encoded in their first language while speaking their second language than in the opposite direction. It seems, therefore, that experience might attenuate the detrimental effect of mismatched linguistic contexts on the speed with which bilinguals access information.

The asymmetry in linguistic experience with mismatched encoding and retrieval contexts underscores, once again, the importance of taking into account language history when conducting and interpreting research with bilinguals. Nevertheless, note that if reaction times reflected only differences in bilinguals' proficiency speaking Spanish and English, we would predict the English–English condition to be much slower for the unbalanced bilinguals. That bilinguals responded faster in the English–English condition, despite differences in proficiency in the two languages in the unbalanced bilinguals, suggests that a match in encoding and retrieval languages does improve memory access and that language-dependent memory effects interact with proficiency and experience effects (and may occasionally override them). We focus the rest of the discussion on potential mechanisms that may be driving language-dependent memory effects.

Mechanisms of language-dependent memory

Language-dependent memory phenomena are situated within the field of context-dependent memory, and follow the same principles as other contexts subsumed by the encoding specificity principle. However, language-dependent memory may also rely on cognitive mechanisms that are especially linked with language. For instance, Slobin (2003, p.177) noted that 'it is quite likely that the language in which information is presented . . . plays a role in the ways in which information is stored and evaluated'. Slobin also suggested that listeners and observers encode information so that it is easy to express linguistically. Encoding information in a way that will be consistent with subsequent retrieval can be formulated as a 'thinking for potential speaking' extension of Slobin's 'thinking for speaking' hypothesis (Slobin, 2003). When applied to bilingual speakers, language-dependent retrieval may be related to bilinguals' expectations about the match and mismatch in the encoding and retrieval languages. Specifically, bilinguals are aware of which languages they speak in which contexts, and may often have well-defined expectations for when they will have to speak each language. Slobin's 'thinking for potential speaking' offers an elegant theoretical framework in which to situate these experiences of bilinguals. If bilinguals have expectations (which need not be explicit) about when linguistic encoding and retrieval contexts will match, they might attend more heavily, or exclusively, to those linguistic distinctions important to the encoding language rather than those of their other language. If bilinguals have expectations about when linguistic encoding and retrieval contexts will not match, they might attend to *all* aspects of a scene that may be required for later linguistic retelling. That is, if bilinguals expect to have to deal with a linguistic context at the time of retrieval that is different from the linguistic context at the time of encoding, they may adjust their encoding strategies and encode information consistent with the language of retrieval, in addition to information consistent with the language of encoding. One outcome of this strategy may be the advantage experienced by bilinguals in contexts such as academic environments, where they are likely to have expectations about matching linguistic environments at learning and testing, when retrieving information in the same language in which it was encoded. Note that if language-dependent memory patterns were 'just' encoding specificity effects, we would not expect the patterns to be modulated by factors such as language proficiency and history of language use.

In addition to the processing considerations of 'thinking for speaking', another potential mechanism for language-dependent memory may be tied to mental representations underlying linguistic expression, and rely on principles of linguistic relativity. Studies in linguistic relativity suggest that language influences how speakers and listeners represent the world (e.g., Boroditsky et al., 2002; 2003; Lucy, 1992; Pederson et al., 1998). Within the framework of linguistic relativity, the content of memory may be affected by the language used to encode memories. Language may be influencing how people remember events by imposing a specific worldview through the linguistic structure and cognitive dimensions it grammaticizes. Levinson (2003) writes, '... given the architecture of the [cognitive] system, once one puts serious semantic constraints on the output, the rest of the system will be forced to support, code, and operate on those features' (p. 301). For bilingual speakers, it may be the case that a match of linguistic encoding and retrieval contexts facilitates recall of language-specific *representations*. These language-specific representations may act as an additional mechanism driving language-dependent memory. Spelke & Tsivkin (2001) provide evidence for this mechanism in the representation and learning of exact numbers. In short, language is likely to serve as a mental frame, aiding the recall and evaluation of representations, as well as a specific cue for encoded information (Marian & Kaushanskaya, in press). Most likely, a combination of mechanisms—encoding specificity, encoding strategy adjustments in 'thinking-for-potential-speaking', language-specific representations, language as a mental frame and a cue—actively interact to yield language-dependent memory effects in bilingual learning, with more experimentation needed to understand their individual and combined influences.

Role of language proficiency

In contrast to balanced bilinguals, who showed consistent language-dependent memory patterns for response accuracy, latency, and error rates, the pattern of language-dependent memory was less stable in unbalanced bilinguals. Unbalanced bilinguals showed language-dependent memory in reaction time analyses, but did not show patterns of language-dependent memory in accuracy analyses. In error analyses, unbalanced bilinguals showed language-dependent memory for some error types (e.g. errors of commission), but not for others (e.g. errors of omission and partial errors).

If these are bona fide differences between groups, one possible explanation may lie in the salience of language. Cognitive processes are likely to differ when using a higher-proficiency language than when using a lower-proficiency language[6] and may result in different salience levels for language as a cue in balanced and unbalanced bilinguals. Specifically, for unbalanced bilinguals, many differences may characterize processing information in one language versus the other, such as greater cognitive load and demands on memory in the less proficient language. Language per se is only one of many variables that change and as a result, it competes with other changes and its salience as a cue may be lower. For balanced bilinguals, however, similar proficiency levels in the two languages may result in highly similar cognitive processing across languages. In the absence of any other differences (e.g. cognitive resources, processing strategy and efficiency, degree of

[6]Indeed, the finding that unbalanced bilinguals performed worse than balanced bilinguals in the Spanish encoding-Spanish retrieval condition suggests that the unbalanced bilinguals experienced greater overall cognitive demands than did the balanced bilinguals. It is likely that the need to use a lower-proficiency language made the experiment as a whole more challenging for unbalanced bilinguals, thus influencing available cognitive resources and resulting in processing differences.

practice using a language), the only difference becomes language. If language proficiency and familiarity increase the strength of linguistic cuing in language-dependent memory, then it is not surprising that those with different proficiency levels in the two languages rely less on linguistic cues and are less susceptible to their influences than balanced bilinguals. Finally, other differences between the two groups, such as current exposure to the two languages (see Table 1), may have further influenced results, with more research necessary to examine how different aspects of a bilingual's linguistic profile influence language-dependent memory patterns.

A possible alternative explanation is that unbalanced bilinguals, who by definition have lower proficiency in one of their languages, may rely more heavily on the higher-proficiency language when encoding information presented in the lower-proficiency language. So although the language of external input is L2, their internal language of encoding may be L1 or a mixture of L1 and L2. Previous research on the effect of language on bilingual memory access found that internal language can differ from external language and patterns of bilingual memory accessibility can differ when compared across external languages versus internal languages of retrieval (e.g. Schrauf & Rubin, 2000). Therefore, it is possible that the language-dependent memory effect in unbalanced bilinguals was attenuated by the fact that the language of internal encoding differed from the language of external input, influencing whether or not there was an actual match or mismatch between encoding and retrieval. Future research will need to test these possible explanations and to further explore the interaction between proficiency and mechanisms of language-dependent memory.

Future directions

It is possible that language-dependent memory manifests itself differently for different types of information. Recent evidence suggests that bilinguals' memory for proper names may be superior to memory for other types of information (e.g. Gollan, Bonanni, & Montoya, 2005). Pilot investigations of memory for names and other types of lexical material are currently under way. We leave the scrutiny of differences in language-dependent memory patterns for different types of information to future research.

Because the ultimate goal of learning is to transfer knowledge into long-term semantic storage, repeated exposure to material and time for memories to consolidate are necessary for long-term learning and should be a focus of future work. Future experiments of language-dependent memory could vary the length of time between learning and testing, the surface form of material and of questions probing learning, and could incorporate rehearsal components, as well as cross-modal learning and testing (e.g. written language). These variables are especially important to consider given the consequences that language-dependent memory may have in applied contexts. That is, in addition to expanding our theoretical knowledge about the relationship between language and memory, this research highlights the importance of sensitivity to linguistic contexts in educational and clinical settings. Language-dependent memory (as measured by both accuracy and reaction times) may serve as a partial explanation for the lower academic achievement reported for native Spanish speakers tested in US public schools (e.g. Coltrane, 2002; Llagas, 2003), who are likely to have encoded at least some memories in their first language that must be retrieved in the academic context of the second language. Consequently, material tested in the classroom in the second language may rely heavily on fundamentals learned outside the classroom in the first language. We reiterate that care should be taken to include awareness of the effects of test language on performance when teachers and clinicians assess cognitive abilities of bilinguals.

ACKNOWLEDGEMENTS

This work was supported in part by Grants NICHD 1R03HD046952-01A1 and NSF BCS-0418495 to the first author. We thank Henrike Blumenfeld and Margarita Kaushanskaya for helpful discussions of this work and Ann Bradlow, Valerie Burt, Nadia Cone, Juan Pedro García, Phillip Goyeneche, Nicole Kalogeropoulos, Alexis Little, Naveen Malik, Cindy Ruff and Li Sheng for their contributions to this project. We appreciate the helpful comments and feedback that Robert F. Belli, Jean Saint-Aubin and three anonymous reviewers provided on an earlier version of this manuscript.

REFERENCES

Barber, S. (2001). *Martha B. Rabbit Jigsaw Book*. China: Barnes and Noble.

Bernardo, A. B. I. (1998). Language format and analogical transfer among bilingual problem solvers in the Phillipines. *International Journal of Psychology, 33*(1), 33–44.

Billman, D., & Krych, M. (1998). Path and manner verbs in action: effects of 'skipping' or 'exiting' on event memory. *Proceedings of the 20th annual meeting of the cognitive science society.* Mahwah, NJ: Lawrence Erlbaum.

Billman, D., Swilley, A., & Krych, M. (2000). Path and manner priming: verb production and event recognition. *Proceedings of the 22nd annual meeting of the cognitive science society.* Mahwah, NJ: Lawrence Erlbaum.

Boroditsky, L., Ham, W., & Ramscar, M. (2002). What is universal in event perception?: comparing English and Indonesian speakers. *Proceedings of the 24th Annual meeting of the cognitive science society.* Mahwah, NJ: Lawrence Erlbaum.

Boroditsky, L., Schmidt, L. A., & Phillips, W. (2003). Sex, syntax, and semantics. In D. Gentner, & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp.61–79). Cambridge, MA: MIT Press.

Bradlow, A. R., Nygaard, L. C., & Pisoni, D. B. (1999). Effects of talker, rate and amplitude variation on recognition memory for spoken words. *Perception and Psychophysics, 61*(2), 206–219.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Coltrane, B. (2002). English language learners and high stakes tests: an overview of the issues. *ERIC Digest*. Retrieved October 1, 2003, from htttp://www.cal.org.ericcll/ DIGEST.

Dalrymple-Alford, E. C., & Aamiry, A. (1969). Language and category clustering in billingual free recall. *Journal of Verbal Learning and Verbal Behavior, 8*, 762–768.

Davies, G. M., & Thomson, D. M. (Eds.). (1988). *Memory in context: Context in memory.* Chichester, UK: Wiley & Sons.

Durgunoglu, A. Y., & Roediger, H. L. (1987). Test differences in accessing bilingual memory. *Journal of Memory and Language, 26*(4), 377–391.

Ervin, S. M. (1961). Learning and recall in bilinguals. *American Journal of Psychology, 74*, 446–451.

Feist, M. I., & Gentner, D. (in press). Spatial language influences memory for spatial scenes. *Memory and Cognition*.

Francis, W. S. (1999a). Analogical transfer of problem solutions within and between languages in Spanish-English bilinguals. *Journal of Memory and Language, 40*(3), 301–329.

Francis, W. S. (1999b). Cognitive integration of language and memory in bilinguals: semantic representation. *Psychological Bulletin, 125*(2), 193–222.

Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition, 83*(1), 49–79.

Gollan, T. H., Bonanni, M. P., & Montoya, R. I. (2005). Proper names get stuck on bilingual and monolingual speakers' tip of the tongue equally often. *Neuropsychology, 19*, 278–287.

Grosjean, F. (1997). Processing mixed languages: issues, Findings, and Models. In A. M. B. de Groot, & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives*. Mahwah, NJ: Lawrence Erlbaum.

Heredia, R., & McLaughlin, B. (1992). Bilingual memory revisited. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 91–103). Amsterdam: Elsevier Science.

Hespos, S. J., & Spelke, E. S. (2004). Conceptual precursors to language. *Nature, 430*, 453–455.
Kintsch, W. (1970). Recognition memory in bilingual subjects. *Journal of Verbal Learning and Verbal Behavior, 9*(4), 405–409.
Kroll, J., & 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. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 145–168). Mahwah, New Jersey: Lawrence 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 and Language, 33*, 149–174.
Lambert, W. E., Havelka, J., & Crosby, C. (1958). The influence of language-acquisition contexts on bilingualism. *Journal of Abnormal and Social Psychology, 56*, 239–244.
Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity.* Cambridge: Cambridge University Press.
Llagas, C. (2003). *Status and trends in the education of Hispanics*. Retrieved October 1, 2003, from http://nces.ed.gov/pubs2003/hispanics.
Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: an example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13*(5), 585–589.
Lucy, J. (1992). *Grammatical categories and cognition: A case study of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.
MacLeod, C. M. (1976). Bilingual episodic memory: acquisition and forgetting. *Journal of Verbal Learning and Verbal Behavior, 15*(4), 347–364.
Marian, V. (2006). Bilingual research methods. In J. Altarriba, & R. R. Heredia (Eds.), *An Introduction to Bilingualism: Principles and processes*. (chapter 2). Mahwah, NJ: Laurence Erlbaum.
Marian, V., Blumenfeld, H., & Kaushanskaya, M. (2005). Comparing self-reported measures of bilinguals' linguistic profile against standardized language assessment tools. Paper presented at the *Tenth Annual Communication Sciences and Disorders Research Forum*, Evanston, IL Manuscript in preparation.
Marian, V., & Kaushanskaya, M. (2004). Self-construal and emotion in bicultural bilinguals. *Journal of Memory and Language, 51*(2), 190–201.
Marian, V., & Kaushanskaya, M. (2005). Language interaction in bilinguals. *Proceedings of the fifth international symposium on bilingualism*, Barcelona, Spain.
Marian, V., & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General, 129*(3), 361–368.
Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology, 57*(3), 203–220.
Nott, C. R., & Lambert, W. E. (1968). Free recall of bilinguals. *Journal of Verbal Learning and Verbal Behavior, 7*, 1065–1071.
Nygaard, L. C., Burt, S. A., & Queen, J. S. (2000). Surface form typicality and asymmetric transfer in episodic memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 26*(5), 1228–1244.
Otheguy, R. & García, O. (1993). Convergent conceptualizations as predictors of degree of contact in US. Spanish. In A. Roca, & J. Lipski (Eds.), *Spanish in the United States: Linguistic contact and diversity* (pp. 135–154). New York: Mouton de Gruyter.
Pederson, E., Danziger, E., Wilkins, D., Levinson, S., Kita, S., & Senft, G. (1998). Semantic typology and spatial conceptualization. *Language, 74*, 557–589.
Peynircioglu, Z. F., & Durgunoglu, A. Y. (1993). Effects of a bilingual context on memory performance. In J. Altarriba (Ed.), *Cognition and culture: A cross-cultural approach to psychology* (pp. 57–75). Amsterdam: Elsevier.
Pinker, S. (1994). *The language instinct*. New York: HarperCollins.
Schrauf, R. W., & Rubin, D. C. (2000). Internal languages of retrieval: the bilingual encoding of memories for the personal past. *Memory and Cognition, 28*, 616–623.
Singh, L., Morgan, J. L., & White, K. S. (2004). Preference and processing: the role of speech affect in early spoken word recognition. *Journal of Memory and Language, 51*(2), 173–189.
Slobin, D. I. (2003). Language and thought online: cognitive consequences of linguistic relativity. In D. Gentner, & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 157–191). Cambridge, MA: MIT Press.

Smith, M. C. (1991). On the recruitment of semantic information for word fragment completion: Evidence from bilingual priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17*, 234–244.
Spelke, E., & Tsivkin, S. (2001). Language and number: a bilingual training study. *Cognition, 78*, 45–88.
Tulving, E., & Thomson, D. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(3), 352–373.
Watkins, M. J., & Peynircioglu, Z. F. (1983). On the nature of word recall: Evidence for linguistic specificity. *Journal of Verbal Learning and Verbal Behavior, 22*, 385–394.
Whorf, B. L. (1956). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*, J. B. Carroll (Ed.), Cambridge, MA: MIT Press.

APPENDIX

STIMULI, ENGLISH VERSION

MYTHOLOGY STORY

The following story is a famous myth of the Espibi people. All the Espibis know this story well because it is told every year on the holiday that celebrates the beginning of winter. On the morning of the first day of winter, while people are sleeping, a bolt of lightening comes from the sky, accompanied by a strong gust of wind. This bolt of lightening starts the winter season and is unique because it is a special shade of red. This red colours the wind as well and for several hours everything touched by this wind becomes red. This red colour of the wind is recognized by certain animals called Begus. Begus are small animals who cannot be seen by humans and they only live in trees that have leaves. Begus must live among things that will eventually fall but have not fallen yet. They cannot survive when all the leaves of the trees are gone, so they need a signal to leave their homes before all the leaves fall. Their signal to leave is the red wind on the first day of winter. When they see the red wind, they start running to keep up with it. As more and more run, they all become red and run even faster. The red lightening bolt that started winter gives its energy to the wind and the wind gives this energy to the Begus to help them run extremely fast. Running so fast lifts them off the ground and for a special moment a red line can be seen leading to the sky. It is the path of the Begus leaving their homes on earth to build a home in clouds. Clouds are good homes for Begus because, like trees, they are filled with things that will fall but have not fallen yet. As all the Espibi people know, the special thing about winter is snow. They believe that snow is so beautiful because the Begus work hard to cover the land in beauty. They want to make the land beautiful because they are so grateful to have a place to live after they have to leave their trees. To celebrate the beginning of winter, Espibis decorate their home with red clouds. You know its winter when the windows and doors of every house have red clouds on them. Snow and clouds hold a special place in the hearts of Espibi people, and if anyone ever looks to the sky and sees a red cloud, he knows that a Begu was thinking of him that day, and he will be blessed with very good luck.

Mythology Questions[7]

In the story about a myth and celebrations, what is the name of the people who celebrate the beginning of winter? In the story about a myth and celebrations, what is the name of the

[7]In each story, the first two questions probed lexical knowledge (e.g. proper names) and were not included in the analyses.

animals who can see wind on the first day of winter? Why do Espibis know their myth about the beginning of winter so well? How do Espibis celebrate the beginning of winter? What natural event starts winter, according to an Espibi myth? What can Espibis see during a special moment on the first day of winter? During winter, what is a sign of good luck to the Espibis? What do Begus recognize that tells them to leave their trees? Why are trees and clouds good places for Begus to live? What do Begus do when they see the red wind and decide they need to leave their trees? What is the original source of the Begus' energy? Why do Begus make beautiful snow, according to this winter myth?

CHEMISTRY STORY

Recently there has been a lot of news published in scientific journals about a discovery at the University of Aida. In the middle of a project that investigated shapes of elements, two Aida scientists, Tesi and Salis, noticed that sparks and a loud sound were coming from one of the lab tables. They walked over to the table and noticed something unusual in their solution of chemicals. They noticed that a new element had formed and that it was a pyramid. The scientists were puzzled because they had run experiments like this before and had never noticed a pyramidal element. The Aida scientists figured out that one of their absent-minded assistants had accidentally raised the heat on their solution and that this heat caused the sparks and the loud sound. They also figured out that the heat turned their solution into the pyramidal element. After repeating their experiment many times, Tesi and Salis confirmed the existence of a new element and they named it Tesalium, as a combination of their own names. With some additional tests, they learned that Tesalium disappears into the solution a few minutes after the loud sound. They thought that nobody had discovered Tesalium in the past because it only exists as a pyramid for a few minutes. The most interesting aspect of Tesalium is that after it disappears it changes itself to behave almost exactly like the other elements in its solution. It seems that Tesalium can behave like any known element, allowing for the possibility of a new category called hidden elements. This discovery of Tesalium as a hidden element has brought several famous chemists to the University of Aida and many more scientists also want to work on this exciting new project. Scientists would like to run more tests to see if the pyramid shape of Tesalium reappears after it disappears for the first time or if it stays hidden forever. Tesi and Salis have been nominated for the International Prize of Chemistry and the University of Aida has given them money to explore how other hidden elements besides the Tesalium pyramid might take shape under high temperatures.

Chemistry Questions

What is the name of a recently discovered element that has been published in scientific journals? In the story about scientific news, where was a new element discovered? Why is this element named Tesalium? What made the scientists go look at their solution of chemicals? Who did something that changed the solution? What is the name of a new category of elements that Tesalium belongs to? What does Tesalium do after it disappears? In general, what were scientists studying when they discovered the new element? What do the scientists think caused the formation of the new element? What kind of recognition have the scientists received for their discovery? What else do researchers want to know about the new element, or elements like it? How has this discovery affected the University of Aida?

HISTORY STORY

In the 32nd century many different groups of people lived on a faraway continent. Two of the countries on this continent were Mepa and Cuni and these countries were separated by a long and deep valley. This valley was the deepest valley in the whole continent. The midpoint of this valley was called the Zone of Siboma. This zone was a large trading center and many people from each country passed through it every day. This zone was very important to each country because many of their citizens worked there and earned a lot of money for their families. However, at the end of the century, an earthquake destroyed most of the Zone of Siboma. This earthquake destroyed almost all of the areas where people worked and many people lost their jobs. The unemployed citizens of each country pressured their governments to improve the situation. The leader of Mepa wanted both countries to work together to rebuild the zone. If they worked together, they could rebuild faster. However, the Cuni government did not want to work together because they wanted to gain total control of the area. The Cuni leader therefore ordered the Mepas to completely withdraw from the Zone of Siboma, to remove all of their workers, or else, they would attack with military force. The Mepas did not withdraw and decided to fight for their country. Thus, the War of Siboma began and the Mepas fought to protect their parts of the zone. This war lasted two years. It ended when the leader of Mepa was assassinated by two Mepa traitors who were paid off by the Cuni government, so the Cunis won the war. After the war, the entire Zone of Siboma became part of their country. The zone was controlled by the Cunis and Mepa has suffered economically ever since.

History Questions

In the war between two countries, the leader of what country was killed? What is the name of the midpoint of the valley between the two countries? In the story about fighting over territory, what geological feature separated two countries? What did citizens of each country do at the Zone of Siboma? What did the leader of Mepa want the other country to do after the earthquake? How did the Zone of Siboma change after the war? How has the country that lost the war suffered? In the story about fighting over territory, what natural disaster sparked a conflict about an important area between two countries? What happened to citizens of the two countries after the earthquake that made them pressure their governments? After the earthquake, why did not the Cuni government help the Mepas rebuild the destroyed area? What did the Cuni leader order the Mepa government to do? How was the leader of Mepa killed?

BIOLOGY STORY

The following story is about some rare types of plants. On the island Fimo, certain plants grow only during its special hot season. The plants grow because of a special combination of soil and rain that are found only on Fimo. The people of Fimo know their plants well because some are deadly but some are very good sources of nutrients. For example, a very popular dessert, in fact, the national dessert, is made from a plant called Mugal. Mugal plants are healthy and safe. Another plant, which looks very similar to a Mugal, however, is dangerous. This plant causes severe headaches that lead to death in people who touch it. Therefore, it is very important for the people of Fimo to understand the specific differences among their special plants. There are some simple rules to help them learn the safe plants and the dangerous plants. First, the plants that grow during the hot season are all flowered. None of the other plants that grow year-round on the island have flowers. So, you only need to worry about plants with flowers. Second, you need to look at the leaves of plants. Plants with spiny leaves are all safe. Nothing else matters about the plants with spiny leaves. Spiny leaves signal a safe plant. The third rule is the most important and called the rule of colour. This rule applies only to plants whose leaves are smooth. Some plants with smooth leaves are safe but some are dangerous. On the smooth-leaved plants, it is important to notice the colour of the flowers. A smooth plant is safe to eat if it has white or yellow flowers. These colours are valued on the island Fimo because plants with these colours are used in many foods and drinks. The colours are so valued that even the Fimo flag is white and yellow. The important Mugal plants have smooth leaves and white flowers. Mugal is used in the national dessert of Fimo. Plants that have smooth leaves and blue flowers are very dangerous. Plants with blue flowers should never be eaten because they cause vomiting. People who live on Fimo know they must pay attention to the leaves and flowers of the plants during the hot season. They know the rules very well, enjoy what their special plants offer them, and remain safe from the deadly plants.

Biology Questions

What is the name of the island where rare types of plants grow for part of the year? What is the name of the plant that is safe and used in the national dessert of the island? Why do rare types of plants grow on Fimo, and nowhere else on Earth, during its hot season? To distinguish safe plants from dangerous plants on Fimo, what two parts of plants do you need to look at? Certain plants are always safe, regardless of the colour of their flowers. What type of plants are these? Why is the Fimo flag white and yellow? What is a non-fatal sickness caused by dangerous plants on Fimo? In the story about some rare plants on an island, why must the people of that island understand differences among their special plants? A dangerous plant looks similar to the Mugal. What does this dangerous plant cause? What part of a plant lets the people of the island know whether it grows in the hot season or not? On what type of plants is the colour of flowers important to notice? What colour flower signals a dangerous plant on the island?