Bilingual recognition memory: Stronger performance but weaker levels-of-processing effects in the less fluent language

Wendy S. Francis · Marisela Gutiérrez

Abstract The effects of bilingual proficiency on recognition memory were examined in an experiment with Spanish–English bilinguals. Participants learned lists of words in English and Spanish under shallow- and deep-encoding conditions. Overall, hit rates were higher, discrimination greater, and response times shorter in the nondominant language, consistent with effects previously observed for lower frequency words. Levels-of-processing effects in hit rates, discrimination, and response time were stronger in the dominant language. Specifically, with shallow encoding, the advantage for the nondominant language was larger than with deep encoding. The results support the idea that memory performance in the nondominant language is impacted by both the greater demand for cognitive resources and the lower familiarity of the words.

Keywords Recognition · Bilingualism · Levels of processing · Word frequency

Bilingual individuals often report difficulty encoding and remembering information in their less fluent language (L2), relative to their more fluent language (L1). Supporting this intuition, previous research has provided some evidence that bilingual recall is worse in L2 (e.g., Durgunoğlu & Roediger, 1987; Glanzer & Duarte, 1971; Lopez & Young, 1974; Nott & Lambert, 1968). In contrast, repetition-priming results have shown that, in some paradigms, bilinguals show a greater benefit from experimental exposures in L2 (Francis, Augustini, & Sáenz, 2003; Francis & Gallard, 2005; Francis

W. S. Francis (✉) · M. Gutiérrez
Department of Psychology, University of Texas at El Paso,
El Paso, TX 79968, USA
e-mail: wfrancis@utep.edu

& Goldmann, 2011). However, it remains unknown how language proficiency affects recognition memory. Recognition memory has been examined in two bilingual studies. The first did not report L1 and L2 data separately (Kintsch, 1970), and the second had only one test language (Durgunoğlu & Roediger, 1987). Therefore, little is known about relative recognition performance in L1 and L2. The present study investigated the effects of bilingual proficiency on recognition memory. As is explained in the following paragraphs, different hypotheses about bilingual recognition memory performance in L2, relative to performance in L1, can be derived on the basis of the greater cognitive load associated with processing L2 words or on the basis of the lower familiarity of L2 words relative to L1 words. These two conceptualizations make opposite predictions.

Bilingual memory processes

Despite calculations suggesting that more than half of the world's population is bilingual (Harris & McGhee-Nelson, 1992), few models of memory have incorporated bilingualism or bilingual proficiency constructs.[1] The most studied issue in bilingual memory has been whether memory representations depend on the original encoding language. Research in this area indicates that translation equivalents of concrete nouns access common conceptual representations in episodic and semantic memory (see Francis, 1999, 2005, for extensive reviews). However, the efficiency and completeness of L1 and L2 encoding may not be equivalent. Since comparisons

---

[1] A notable exception was Paivio's dual-coding theory, which he extended to bilingualism by explicitly incorporating separate verbal systems for each language and a common image system through which concrete translation equivalents could share a common referent (Paivio & Desrochers, 1980).

of L1 and L2 memory performance were not a major focus of past research, many studies had "balanced" bilingual participants who would be unlikely to show language differences. Also, most studies did not have full factorial designs; some used only one test language, and others included only between-language conditions. In some cases, L1 and L2 data were simply not reported separately. Thus, few studies tested and reported within-language memory performance in each language for bilinguals who had a clear L1, and these studies did not focus on recognition memory.

The present research focuses on understanding the mechanisms by which bilingual language proficiency impacts memory processes. In the absence of an established theory of bilingual proficiency and memory, we use ideas from other domains of bilingual research and from other domains of memory research to make predictions about bilingual memory. In other domains of bilingual research, performance differences for more and less fluent languages have been explained either by the greater demand that L2 processing puts on cognitive resources or by the lower familiarity of L2 words. The effects of cognitive resource availability and familiarity have been extensively studied in monolingual memory research by using divided attention and word frequency manipulations. Therefore, the established effects of divided attention and word frequency on memory were used to make predictions about language proficiency effects on recognition memory. It should be noted that these two explanations are not mutually exclusive and that they lead to common predictions for some memory performance measures; for example, both lead to the prediction of lower free recall performance in L2 than in L1.

L2 processing as working with limited resources Working in L2 makes greater demands on cognitive resources for attention and working memory (Abu-Rabia, 2003; Ransdell, Arecco, & Levy, 2001; Takano & Noda, 1993) and reduces the amount of information that can be held in working memory (e.g., da Costa Pinto, 1991; Service, Simola, Metsaenheimo, & Maury, 2002). Given that processing L2 requires more cognitive resources than does processing L1, working in L2 would be expected to affect memory in the same manner that resource limitations affect memory performance in L1.

Effects of resource limitations on memory have been explored primarily using divided attention manipulations. Attention manipulations are accomplished using dual-task procedures, in which participants perform a secondary task at the same time as the primary task. Adverse effects have been found for both recall (Craik, Naveh-Benjamin, Ishaik, & Anderson, 2000; Naveh-Benjamin, Craik, Perretta, & Tonev, 2000; Whiting, 2003) and recognition memory (Hicks

& Marsh, 2000; Naveh-Benjamin, Craik, Guez, & Dori, 1998). However, these effects have tended to be larger for tasks like recall that rely on associative interitem processing and smaller effects on tasks like recognition that rely primarily on intrinsic item information (Troyer & Craik, 2000), suggesting that divided attention interrupts associative processing. The adverse effects of divided attention on memory performance may be explained in part by reducing the likelihood or effectiveness of memory-enhancing procedures, such as elaboration (Craik & Kester, 2000; Naveh-Benjamin et al., 2000). Similarly, the greater attentional demands necessary to process L2 words, relative to L1 words, may reduce the likelihood or efficacy of elaboration and other procedures that might enhance recognition memory. Therefore, this cognitive resource conceptualization suggests that recognition performance will be worse in L2 than in L1.

L2 processing as working with unfamiliar vocabulary Words in L2 are less familiar and have occurred less often in a person's lifetime than words in L1. Existing theories of bilingual lexical processing, including the revised hierarchical model (Kroll & Stewart, 1994), the bilingual interactive activation model (Dijkstra & Van Heuven, 2002), and the inhibitory control model (Green, 1998), all include the feature that associations between words and their meanings are weaker in L2 than in L1. A *weaker links* hypothesis has been proposed as a single mechanism for bilingual proficiency effects and word frequency effects (Gollan, Montoya, Cera, & Sandoval, 2008). Bilinguals may therefore process L2 words much as they process low-frequency (LF) words in L1 (e.g., Ardila, 2003; Gollan et al., 2008). Both LF and L2 words are at an earlier point on a learning curve and are more weakly associated with their meanings than are high-frequency (HF) and L1 words (Gollan et al., 2008). Experiences with both LF and L2 words may be limited to fewer contexts, and an LF or L2 word may have greater orthographic distinctiveness. A factor that differentiates L2 words from LF words in L1 is that L2 words may be associated with concepts that already have strong links formed within L1. Nevertheless, lower familiarity of the L2 verbal label may make it more difficult to integrate L2 content with long-term memory representations.

Although recall exhibits an advantage for HF words, LF words are better recognized (Balota & Neely, 1980; Dewhurst, Hitch, & Barry, 1998; Kinsbourne & George, 1974; MacLeod & Kampe, 1996; Mandler, Goodman, & Wilkes-Gibbs, 1982), exhibiting the *mirror effect*: both higher hit rates and lower false alarm rates. The same pattern is observed for words that are normatively learned early or late, with later-learned words being better recognized (Dewhurst et al., 1998). According to Reder's *source-of-activation confusion* explanation (e.g., Diana & Reder,

2006), the mirror effect arises because LF words are preexperimentally associated with fewer episodic contexts, which makes it easier at test to discriminate the experimental episode from previous episodes in which an LF word was encountered. Similarly, L2 words have been associated with fewer episodic contexts than have L1 words; therefore, we might expect L2 words, like LF words, to exhibit better discrimination of the study episode from preexperimental episodes. Therefore, the familiarity conceptualization suggests that recognition performance will be better in L2 than in L1.

Levels-of-processing effects

The resource limitation and familiarity approaches to bilingual memory described above also lead to divergent predictions about the effects of encoding manipulations on L1 and L2 recognition memory performance. *Levels-of-processing (LOP) effects* refer to the phenomenon that deep conceptual processing at study generally leads to better memory performance than does shallow nonsemantic processing (Craik & Lockhart, 1972). LOP effects are robust in both recall and recognition (e.g., Craik & Lockhart, 1972). The mechanism for these effects is thought to be that elaboration makes it easier to differentiate the memory of an experimental episode of studying an item from other memory episodes (Craik, 2002). Divided attention reduces LOP effects in recognition by differentially hurting deeply processed items (Hicks & Marsh, 2000). In contrast, LF words benefit more from deep encoding than do HF words in both speed and accuracy (Duchek & Neely, 1989). LOP effects have not been compared across languages in bilingual recall or recognition.

The present study

The present study compares L1 and L2 recognition memory performance. Divided attention and word frequency approaches to L2 processing lead to contrasting predictions for accuracy of bilingual recognition performance and contrasting predictions about the effects of an LOP manipulation. The negative effect of division of attention on recognition memory performance suggests that performance will be worse in L2 than in L1 recognition. The weakening of LOP effects under divided attention suggests that LOP effects will be weaker in L2 than in L1. In contrast, the advantage for LF words in recognition memory suggests that recognition will be better in L2 than in L1. Also, the larger LOP effects observed for LF over HF words suggests that LOP effects will be stronger in L2 than in L1.

These resource limitation and unfamiliar vocabulary conceptualizations of L2 processing are not necessarily mutually exclusive. In fact, in monolinguals, there is an interaction between divided attention and word frequency effects on recognition, with divided attention being more detrimental to LF than to HF words (Diana & Reder, 2006). The effects of a combination of resource limitation and familiarity factors on bilingual recognition can be derived.

First, applying the word frequency conceptualization of L2 processing and the source-of-activation confusion account, L2 words would exhibit better recognition performance than L1 words because, in L2, the experimental study episode will be more discriminable from preexperimental episodes. Second, associations between all words and their experimental study episodes become stronger under deep processing than under shallow processing (Craik, 2002), thus producing the LOP effect. Finally, we consider the finding that under resource limitation, divided attention hurts performance by differentially hurting deeply processed items (Hicks & Marsh, 2000). Applying the expected divided attention effects to L2 words only, deeply processed words would be hurt more by being studied in L2 than would shallowly processed words, thus creating an interaction with a smaller LOP effect in L2 and a smaller language difference for deeply processed items.

The research question is whether the cognitive load conceptualization, the word familiarity conceptualization, or a hybrid conceptualization would better explain bilingual recognition performance in L2 and L1.

Method

Participants

Participants were 64 self-identified Spanish–English bilinguals (24 men, 40 women), who participated to fulfill a research requirement for an introductory psychology class at the University of Texas at El Paso. They were recruited through an electronic sign-up system. They ranged in age from 17 to 41 years (median = 19), and all but one reported Hispanic ethnicity. On the basis of self-reported relative proficiency, 50% were classified as English dominant, and 50% were classified as Spanish dominant. The first language learned was reported as Spanish for 88% and English for 6%; 6% reported having learned both languages simultaneously from early childhood. The median age of L2 acquisition was 6 years. Usage over the preceding month was reported as 46% English, 45% Spanish, 9% a mixture of English and Spanish, and less than 1% other languages; this pattern corresponded to using the dominant language 61% of the time and the nondominant language 30% of the time. Most of the participants were residents of El Paso, Texas, which is an English–Spanish double-

immersion environment, even at the university. A number of the Spanish-dominant participants were students from the adjacent Ciudad Juárez, Chihuahua, in Mexico, who commuted across the international border daily to attend class. Four additional individuals completed the protocol but were excluded, 3because of insufficient proficiency in Spanish and 1 because of a form error.

Design

The experimental conditions formed a 2 (dominant language) × 2 (task language) × 2 (encoding condition) mixed design. Half of the participants were English dominant, and half were Spanish dominant. Each participant studied one list in English and one in Spanish. Half of the items in each language were studied, of which half were processed using a shallow-encoding task and half were processed using a deep-encoding task.

Materials and apparatus

The stimuli were 216 concrete nouns in English and Spanish, chosen to be relatively unambiguous in meaning and likely to be in the vocabulary of the participants. The mean letter lengths in English and Spanish were 5.3 and 6.2, respectively. Their median frequency in the language was 15 per million in English (Kučera & Francis, 1967) and 13 per million in Spanish (Alameda & Cuetos, 1995). The words were randomly assigned to eight sets of 27 words. Half of the sets were assigned to each language; within each language, one set was assigned to shallow processing, one was assigned to deep processing, and two were assigned to be foil items in the recognition test. The sets were rotated through language and encoding conditions across participants, using a Latin square to control for specific-item effects.

The stimuli were presented on the monitor of a Macintosh computer. The sequence of stimulus presentation and timing of responses was programmed using PsyScope software (Cohen, MacWhinney, Flatt, & Provost, 1993). Recognition responses were collected using a PsyScope button box (New Micros, Dallas).

Procedure

Participants were tested individually by a bilingual experimenter in sessions lasting approximately 45 min. Instructions were given in the language of each task. Participants learned and recognized one set of words in English and one set of words in Spanish, with the language order counterbalanced across participants. The study and recognition tasks were computerized. Each study sequence included 54 words. Half of the experimental words were processed under shallow-

encoding instructions, and half under deep-encoding instructions, randomly intermixed. The shallow task was to indicate the number of vowels in each word by pressing the appropriate number on the keyboard. The deep task was to determine whether the object that the word referred to was natural or manufactured and to press the n or m key to indicate the response. On each trial, a cue appeared 1.5 s before each word to alert the participant to which task should be performed and remained on the screen with the word until a response was registered. After an intertrial interval of 500 ms, the next cue appeared.

The recognition test immediately followed completion of the study sequence. At test, the 54 studied words and 54 nonstudied words were presented one at a time in a randomly intermixed order. Participants were to indicate as quickly and accurately as possible whether each word was studied or not by pressing a yes or no key on the button box. To reduce interference, upon completion of the first study–test sequence, at least 10 min intervened before initiation of the second study–test sequence in the other language. During this interval, participants completed a language background questionnaire and then tried to solve a Rubik's cube until 10 min had passed.

Results

Hits and false alarms

Hit rates and false alarm rates are shown in Table 1. For inferential analysis, these values were arcsine transformed. Transformed hit rates for studied items were submitted to a 2 (dominant language) × 2 (task language) × 2 (LOP) mixed ANOVA. The main effects of dominant language and task language did not approach significance, $Fs < 1$. However, a significant interaction of these factors, $F(1, 62) = 9.63$, $MSE = .0257$, $p = .003$, indicated that hit rates were higher in the nondominant language than in the dominant language. As in previous research, a significant LOP effect, $F(1, 62) = 312.00$, $MSE = .0302$, $p < .001$, showed higher hit rates for words in the deep-encoding condition than for words in the shallow-encoding condition. The LOP effect did not enter into two-way interactions with dominant language or task language, $Fs < 1$. However, the three-way interaction was significant, $F(1, 62) = 5.11$, $MSE = .0262$, $p = .027$, indicating that the LOP effect on hit rates was stronger in the dominant language than in the nondominant language. The three-way interaction also indicated that the advantage for the nondominant language was reliable in the shallow-encoding condition, $F(1, 62) = 16.03$, $MSE = .0232$, $p < .001$, but not in the deep-encoding condition, $F < 1$. False alarm rates were submitted to a 2 (dominant language) × 2 (task language) mixed ANOVA. The main effects of

dominant language and task language and their interaction did not approach significance, Fs < 1.

Because several participants had either 100% hits or 0% false alarms in at least one condition, d' could not be computed. Therefore, the A' statistic was used to estimate discrimination of studied from nonstudied items. A' was calculated for each condition and is listed in Table 1. Because false alarm rates did not differ across languages, the A' analysis was redundant with the hit rate analysis, with the same pattern of significant effects.

Response times

For the response time (RT) analysis, only correct responses were included, and trials with RTs greater than 3,000 ms or less than 300 ms were excluded as outliers. RTs to correctly recognized presented words were submitted to a 2 (dominant language) × 2 (task language) × 2 (LOP) mixed ANOVA. The main effects of dominant language, F < 1, and task language, $F(1, 62) = 1.198$, $MSE = 22,356$, $p = .278$, did not approach significance. However, a significant interaction of these factors, $F(1, 62) = 33.005$, $MSE = 22,356$, $p < .001$, showed that recognition responses were faster in the less fluent language. As in previous research,

RTs exhibited an LOP effect, $F(1, 62) = 102.441$, $MSE = 18,406$, $p < .001$, with faster responses for words processed under deep encoding than for words processed under shallow encoding. LOP did not enter into two-way interactions with dominant language or task language, Fs < 1. However, a significant three-way interaction, $F(1, 62) = 12.954$, $MSE = 12,179$, $p = .001$, indicated a stronger effect of LOP on RTs in the dominant language than in the nondominant language. The three-way interaction also indicated that the L2 advantage was stronger for shallow encoding than for deep encoding, but unlike the hit rate and discrimination analyses, the advantage for the nondominant language was significant for both shallow-encoding, $F(1, 62) = 35.48$, $MSE = 22,239$, $p < .001$, and deep-encoding, $F(1, 62) = 8.67$, $MSE = 12,297$, $p = .005$, conditions.

RTs for correct rejections of foil items were subjected to a 2 (dominant language) × 2 (task language) mixed ANOVA. The main effect of dominant language did not approach significance, $F < 1$. The main effect of task language, $F(1, 62) = 5.306$, $MSE = 11,539$, $p = .025$, indicated faster responses in English than in Spanish. Most important, a significant interaction of dominant language and task language, $F(1, 62) = 17.079$, $MSE = 11,539$, $p < .001$, showed that RTs to reject foil items were faster in the less fluent language.

Table 1 Recognition performance as a function of language dominance, task language, and level of processing at encoding

Language Dominance Task Language Encoding Condition	% Yes Responses	A'	RT Correct (ms)
English Dominant ($N = 32$)			
English (L1)			
Shallow	61.1	.934	1,250
Deep	87.0	.993	1,009
LOP effect	(25.9)	(.059)	(241)
Not presented	10.2	–	1,187
Spanish (L2)			
Shallow	67.5	.946	1,110
Deep	87.5	.992	975
LOP effect	(20.0)	(.046)	(135)
Not presented	9.4	–	1,153
Spanish Dominant ($N = 32$)			
English (L2)			
Shallow	68.1	.954	1,050
Deep	86.9	.995	941
LOP effect	(18.8)	(.041)	(109)
Not presented	11.2	–	1,076
Spanish (L1)			
Shallow	59.0	.929	1,224
Deep	86.0	.991	1,023
LOP effect	(27.0)	(.062)	(201)
Not presented	10.2	–	1,199

Discussion

As in previous research, LOP effects were evident in hit rates, discrimination, and RTs, with better performance for words studied under deep-encoding instructions than for words studied under shallow-encoding instructions. There were two new main findings. First, recognition performance was better in the less fluent language in terms of both accuracy and speed. Specifically, L2 performance exhibited higher hit rates, higher discrimination scores, and shorter RTs than did L1 performance. False alarm rates and, therefore, the response criterion were comparable for the two languages; thus, there was no language-based mirror effect. The L2 advantage is consistent with previous research showing an advantage for LF over HF words (e.g., Balota & Neely, 1980; MacLeod & Kampe, 1996) and an advantage for late-acquired over early-acquired words (Dewhurst et al., 1998). Thus, like LF words, L2 words have a disadvantage in recall but an advantage in recognition. This result is inconsistent, however, with the prediction based on divided attention that discrimination would be greater in the more fluent language.

The second main finding was a stronger LOP effect in L1 than in L2, which was evident in hit rates, discrimination, and RTs. A decomposition of this interaction showed similar accuracy and discrimination in the two languages following deep encoding but better L2 than L1 performance following shallow encoding. Note, however, that the L2 advantage in RT was significant for both shallow and deep processing. This finding cannot be explained by either the word frequency or divided attention conceptualizations alone. The weaker LOP effect in L2 is inconsistent with predictions based solely on the lower frequency of L2 words, in that LF words showed a greater LOP effect than did HF words (Duchek & Neely, 1989). The weaker LOP effect in L2 is also inconsistent with predictions based solely on the greater attentional resources needed for L2 processing. Specifically, dividing attention at encoding diminished the LOP effect by differentially hurting deeply encoded words (Hicks & Marsh, 2000). In the present study, it appears, instead, as if shallowly encoded words were helped by being presented in L2. Thus, neither conceptualization alone adequately explains the second main finding of the present study.

A combined conceptualization, as described in the introduction, provides a possible explanation of the pattern of effects observed. This conceptualization correctly predicts that increased discriminability from preexperimental episodes will lead to better recognition performance for L2 than for L1 words. It also correctly predicts the LOP effect, based on strengthened associations between all words and their experimental study episodes for deep, relative to shallow, processing (Craik, 2002). Finally, it correctly

predicts that L2 processing will differentially hurt deeply processed items as does dividing attention (Hicks & Marsh, 2000), thus eliminating the L2 advantage under deep-processing conditions. For shallowly processed words, the L2 advantage was maintained in hit rates, discrimination, and RT, showing that the advantage of greater episodic distinctiveness in L2 outweighed the disadvantage in L2 processing requirements. In contrast, for deeply processed words, the two effects canceled each other out to some degree in hit rates and discrimination, but in RT, the L2 advantage was maintained.

This explanation is consistent with previous research showing that recall performance is particularly dependent on associative processing, whereas recognition performance is based more on individual item processing (Hunt & Einstein, 1981; Mandler, 1980). Because divided attention appears to interrupt associative processing, it tends to have smaller effects on recognition than on recall (Troyer & Craik, 2000). The adverse effects of divided attention on memory performance may be explained in part by reducing the likelihood or effectiveness of memory-enhancing procedures, such as elaboration (Craik & Kester, 2000; Naveh-Benjamin et al., 2000), and therefore, deeply processed items are differentially hurt (Hicks & Marsh, 2000). Thus, we suspect that the greater attentional demands necessary to process L2 words, relative to L1 words, differentially affected the deep-processing condition, where there was more elaborative processing to disrupt.

One alternative explanation based on the present data alone would be that under shallow conditions, L2 words are encoded more richly than L1 words, thus leading to the observed interaction. However, such an explanation cannot account for the previous findings that L2 recall is worse than L1 recall. This is similar to the LF advantage in recognition, which cannot be attributed to deeper processing of LF relative to HF words, because HF words are better recalled.

The present results may, in fact, underestimate the true effects of language proficiency on recognition performance and on levels-of-processing effects. First, a limitation of the study is that language dominance was assessed using self-report measures, rather than with objective language assessments. Therefore, to the extent that some individuals may have been misclassified in terms of language dominance, we may have underestimated the true effects of bilingual proficiency. Second, the bilingual participants were highly proficient in both languages and would be expected to show smaller proficiency effects than would less balanced bilinguals. Finally, the high hit rates in the deep-processing conditions and the low false alarm rates may not have left enough room to detect effects of language proficiency in deep-condition hit rates or in false alarms.

Conclusions

Overall, recognition memory performance was stronger in L2 than in L1, consistent with effects previously observed for lower frequency words. However, this L2 advantage was moderated by an interaction with LOP. The LOP effect was weaker in L2 than in L1, contrasting with the previous finding that LOP effects in recognition were stronger in LF than in HF words (Duchek & Neely, 1989). With shallow encoding, the L2 advantage was larger than with deep encoding. The results support the idea that bilingual memory performance in a less fluent language is impacted by both the greater demand for cognitive resources and the lower familiarity of the L2 words.

References

Abu-Rabia, S. (2003). The influence of working memory on reading and creative writing processes in a second language. *Educational Psychology, 23,* 209–219.

Alameda, J. R., & Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano*. Oviedo: Servicio de Publicaciones de la Universidad de Oviedo.

Ardila, A. (2003). Language representation and working memory with bilinguals. *Journal of Communication Disorders, 36,* 233–240.

Balota, D. A., & Neely, J. H. (1980). Test-expectancy and word-frequency effects in recall and recognition. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 576–587.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers, 25,* 257–271.

Craik, F. I. M. (2002). Levels of processing: Past, present. .. and future? *Memory, 10,* 305–318.

Craik, F. I. M., & Kester, J. D. (2000). Divided attention and memory: Impairment of processing or consolidation? In E. Tulving (Ed.), *Memory, consciousness, and the brain: The Tallinn Conference* (pp. 38–51). Philadelphia, PA: Psychology Press.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior, 11,* 671–684.

Craik, F. I. M., Naveh-Benjamin, M., Ishaik, G., & Anderson, N. D. (2000). Divided attention during encoding and retrieval: Differential control effects? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1744–1749.

Da Costa Pinto, A. (1991). Reading rates and digit span in bilinguals: The superiority of mother tongue. *International Journal of Psychology, 26,* 471–483.

Dewhurst, S. A., Hitch, G. J., & Barry, C. (1998). Separate effects of word frequency and age of acquisition in recognition and recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 284–298.

Diana, R. A., & Reder, L. M. (2006). The low-frequency encoding disadvantage: Word frequency affects processing demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 805–815.

Dijkstra, T., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language & Cognition, 5,* 175–197.

Duchek, J. M., & Neely, J. H. (1989). A dissociative word-frequency× levels-of-processing interaction in episodic recognition and lexical decision tasks. *Memory & Cognition, 17,* 148–162.

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

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

Francis, W. S. (2005). Bilingual semantic and conceptual representation. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 251–267). New York, NY: Oxford University Press.

Francis, W. S., Augustini, B. K., & Sáenz, S. P. (2003). Repetition priming in picture naming and translation depends on shared processes and their difficulty: Evidence from Spanish–English bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 1283–1297.

Francis, W. S., & Gallard, S. L. K. (2005). Concept mediation in trilingual translation: Evidence from response time and repetition priming patterns. *Psychonomic Bulletin & Review, 12,* 1082–1088.

Francis, W. S., & Goldmann, L. L. (2011). Repetition priming within and between languages in semantic classification of concrete and abstract words. *Memory, 19,* 653–663.

Glanzer, M., & Duarte, A. (1971). Repetition between and within languages in free recall. *Journal of Verbal Learning and Verbal Behavior, 10,* 625–630.

Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language, 58,* 787–814.

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

Harris, R. J., & McGhee-Nelson, E. M. (1992). Bilingualism: Not the exception any more. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 3–14). Amsterdam: Elsevier.

Hicks, J. L., & Marsh, R. L. (2000). Toward specifying the attentional demands of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1483–1498.

Hunt, R. R., & Einstein, G. O. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior, 20,* 497–514.

Kinsbourne, M., & George, J. (1974). The mechanism of the word-frequency effect on recognition memory. *Journal of Verbal Learning and Verbal Behavior, 13,* 63–69.

Kintsch, W. (1970). Recognition memory in bilingual subjects. *Journal of Verbal Learning and Verbal Behavior, 9,* 405–409.

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.

Kučera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.

López, M., & Young, R. K. (1974). The linguistic interdependence of bilinguals. *Journal of Experimental Psychology, 102,* 981–983.

MacLeod, C. M., & Kampe, K. E. (1996). Word frequency effects on recall, recognition, and word fragment completion tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 132–142.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87,* 252–271.

Mandler, G., Goodman, G. O., & Wilkes-Gibbs, D. L. (1982). The word-frequency paradox in recognition. *Memory & Cognition, 10,* 33–42.

Naveh-Benjamin, M., Craik, F. I. M., Guez, J., & Dori, H. (1998). Effects of divided attention on encoding and retrieval processes in human memory: Further support for an asymmetry. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1091–1104.

Naveh-Benjamin, M., Craik, F. I. M., Perretta, J. G., & Tonev, S. T. (2000). The effects of divided attention on encoding and retrieval processes: The resiliency of retrieval processes. *Quarterly Journal of Experimental Psychology, 53A,* 609–625.

Nott, C. R., & Lambert, W. E. (1968). Free recall of bilinguals. *Journal of Verbal Learning and Verbal Behavior, 7,* 1065–1071.

Paivio, A., & Desrochers, A. (1980). A dual-coding approach to bilingual memory. *Canadian Journal of Psychology, 34,* 388–399.

Ransdell, S., Arecco, M. R., & Levy, C. M. (2001). Bilingual long-term working memory: The effects of working memory loads on writing quality and fluency. *Applied PsychoLinguistics, 22,* 113–128.

Service, E., Simola, M., Metsaenheimo, O., & Maury, S. (2002). Bilingual working memory span is affected by language skill. *European Journal of Cognitive Psychology, 14,* 383–407.

Takano, Y., & Noda, A. (1993). A temporary decline of thinking ability during foreign language processing. *Journal of Cross-Cultural Psychology, 24,* 445–462.

Troyer, A. K., & Craik, F. I. M. (2000). The effect of divided attention on memory for items and their context. *Canadian Journal of Psychology, 54,* 161–171.

Whiting, W. L. (2003). Adult age differences in divided attention: Effects of elaboration during memory encoding. *Aging, Neuropsychology, and Cognition, 10,* 141–157.