The bilingual L2 advantage in recognition memory

Wendy S. Francis · E. Natalia Strobach

Published online: 19 April 2013
© Psychonomic Society, Inc. 2013

Abstract To better understand the mechanisms by which bilingual proficiency impacts memory processes, two recognition memory experiments were conducted with matched monolingual and bilingual samples. In Experiment 1, monolingual speakers of English and Spanish studied high- and low-frequency words under full attention or cognitive load conditions. In Experiment 2, Spanish–English bilingual participants studied high- and low-frequency words under full-attention conditions in each language. For both monolinguals and bilinguals, low-frequency words were better recognized than high-frequency words. The central new findings were that bilingual recognition was more accurate in the less fluent language (L2) than in the more fluent language (L1) and that bilingual L2 recognition was more accurate than monolingual recognition. The bilingual L2 advantage parallels word frequency effects in recognition and is attributed to the greater episodic distinctiveness of L2 words, relative to L1 words.

Keywords Recognition memory · Bilingualism · Word frequency · Cognitive load

The bilingual L2 advantage in recognition memory

Despite calculations suggesting that over half of the world's population is bilingual (Harris & McGhee-Nelson, 1992), little research has been devoted to understanding the impact of bilingual language proficiency on verbal memory. Although several studies indicate that recall performance is less accurate in a bilingual's less fluent language (e.g., Durgunoglu & Roediger, 1987; Glanzer & Duarte, 1971), the impact of proficiency on other measures of explicit memory is relatively unexplored. The present study focuses on two primary questions. First, does bilingual recognition memory performance differ for the more and less fluent languages?

Second, does recognition performance in bilinguals differ from that of monolinguals?

Given previous recall results, it is tempting to assume that bilingual memory performance will always suffer deficits in the less fluent language (L2), relative to the more fluent language (L1) and relative to monolingual performance. However, predictions of bilingual performance depend on the mechanisms by which language proficiency is thought to affect memory encoding and retrieval. Two aspects of bilingual language processing are of particular relevance to memory. First, L2 processing makes greater demands on cognitive resources than does L1 processing (Abu-Rabia, 2003; Ransdell, Arecco, & Levy, 2001). This demand on cognitive resources is related to the construct of cognitive load, which has known effects on explicit memory performance. The cognitive load imposed by performing a concurrent task at study impairs both recall and recognition (e.g., Hicks & Marsh, 2000; Naveh-Benjamin, Craik, Guez, & Dori, 1998). The mechanism underlying this impairment is thought to be a reduction in the use of elaborative processing (Naveh-Benjamin, Craik, Perretta, & Tonev, 2000). Similarly, the cognitive load associated with processing L2 words may reduce the use of elaborative processing and thereby impair L2 recognition, relative to L1 recognition and monolingual recognition.

The second relevant aspect of bilingual language processing is that L2 words are more weakly associated than L1 words with the concepts that they represent in semantic memory. This feature is incorporated in existing models 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). These associations also appear to be weaker in bilingual L1 words than in monolinguals, and a weaker-links hypothesis has been proposed as a single mechanism for bilingual/monolingual differences, bilingual proficiency effects, and word frequency effects (Gollan, Montoya, Cera, & Sandoval, 2008). Word frequency has opposite effects on recall and recognition. While high-frequency words are better recalled, low-frequency words are

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

better recognized (Balota & Neely, 1980; Kinsbourne & George, 1974; MacLeod & Kampe, 1996; Mandler, Goodman, & Wilkes-Gibbs, 1982), exhibiting the *mirror effect*, both higher hit rates and lower false alarm rates. According to the *source-of-activation-confusion* theory (Buchler & Reder, 2007; Diana & Reder, 2006), the mirror effect arises for two reasons. First, because low-frequency words are preexperimentally associated with fewer episodic contexts, there is less contextual competition, which facilitates recollection and increases hit rates. Second, because low-frequency words have a lower degree of strength in memory, it is less likely that a familiarity-based false alarm will occur. Similarly, L2 words are likely to have occurred in fewer episodic contexts and may therefore exhibit higher hit rates and lower false alarm rates than L1 words and words in a monolingual person's vocabulary.

The cognitive load and word frequency conceptualizations of L2 memory lead to opposite predictions for recognition performance. Specifically, the cognitive load approach suggests that L2 recognition will be worse than L1 recognition, whereas the word frequency approach suggests that L2 recognition will be better than L1 recognition. The purpose of the present study is to determine which of the two conceptualizations better explains L2 performance in the context of recognition memory.

One preliminary study provides some evidence for superior recognition in L2, relative to L1, following incidental encoding (Francis & Gutiérrez, 2012). Specifically, an advantage in hit rates (but not false alarm rates) was observed for words studied under shallow encoding, but not under deep encoding instructions. The effects may have been underestimated because of ceiling and floor effects for hit and false alarm rates and because language dominance was self-reported. In the present study, we gave intentional learning instructions, used a more difficult recognition task, assessed language proficiency and dominance objectively, included matched English and Spanish monolingual comparison groups, and included manipulations of cognitive load and word frequency. The study was conducted on the U.S.–Mexico border at the University of Texas at El Paso (UTEP) and at the neighboring Universidad Autónoma de Ciudad Juárez (UACJ) in Mexico. The city of El Paso and UTEP itself are bilingual environments in which there is ample opportunity to use English and Spanish on a daily basis.

The primary purpose of Experiment 1 was to provide anchor points for monolingual performance in English and Spanish. The secondary purpose was to replicate the effects of cognitive load and word frequency on recognition memory in monolingual English-speaking and monolingual Spanish-speaking participants. That is, we wanted to ensure that the stimuli to be used in Experiment 2 would produce cognitive load and word frequency effects. The purpose of Experiment 2 was to determine whether the cognitive load or word

frequency conceptualization of L2 memory would better explain recognition performance. Experiment 2 tested for the first time the effects of language proficiency and word frequency on recognition memory in bilinguals following intentional encoding. To enable comparisons between experiments and populations, participants in both experiments completed standardized language and cognitive assessments and answered questions about socioeconomic status.

Experiment 1

Experiment 1 was conducted to provide monolingual anchor points for recognition performance and to ensure that the stimuli used would indeed produce cognitive load and word frequency effects as in previous research. In Experiment 1, monolingual speakers of English and Spanish completed two recognition memory study–test cycles. In the full-attention condition, participants read aloud and attempted to memorize high- and low-frequency words. In the cognitive load condition, participants completed the same study task while performing a concurrent *n*-back task. In both conditions, the test task was a yes–no recognition task.

Method

Participants

Participants were 32 monolingual English speakers and 32 monolingual Spanish speakers. The English speakers were students from UTEP, and the Spanish speakers were students from UACJ. All participants were paid. Participant characteristics are summarized in Table 1. Monolingual status was verified using short forms of the Woodcock–Muñoz Language Survey–Revised (WMLS–R) in English and Spanish.

Design

The experimental conditions formed a 2 (language group) $\times$ 2 (attention level) $\times$ 2 (frequency) mixed design. Half of the participants were English speakers, who completed the protocol in English, and half were Spanish speakers, who completed the protocol in Spanish. Attention and word frequency were manipulated within subjects.

Materials

The experimental stimuli were 160 high-frequency words and 160 low-frequency words in English and their translation equivalents in Spanish. High-frequency words had frequencies greater than 75 per million, and low-frequency words had frequencies lower than 10 per million in both languages (Alameda & Cuetos, 1995; Kucera & Francis, 1967). All

Table 1 Characteristics of monolingual and bilingual participants

	Monolingual English (N = 32)	Monolingual Spanish (N = 32)	Bilingual English Dom (N = 32)	Bilingual Spanish Dom (N = 32)
Median age	22	22	22	21
Median age of acquisition for English	1	–	5	7
Median age of acquisition for Spanish	–	1	1	1
Age equivalency English picture vocabulary	25.1	4.3	19.4	10.1
Age equivalency Spanish picture vocabulary	3.3	20.4	11.5	20.6
Age equivalency visual-spatial reasoning ability	20.8	20.6	19.9	19.0
Age equivalency fluid reasoning ability	17.6	17.2	17.8	17.8
Median of highest parental education level	Some college	Graduated college	Some college	Graduated college

Dom = dominant

selected words had relatively unambiguous translation equivalents. The high- and low-frequency words were randomly divided into four subsets matched on frequency and letter length. The assignment of lists to full-attention and cognitive load conditions and to studied or foil status was counterbalanced across participants.

Procedure

Participants completed the required tasks in two sessions. In session 1, the short version of the WMLS–R (Woodcock, Muñoz-Sandoval, Ruef, & Alvarado, 2005) was administered in English and Spanish by a bilingual experimenter. Selected nonverbal tests from the Woodcock–Johnson Tests of Cognitive Abilities (WJ–III; Woodcock, McGrew, & Mather, 2001) were administered to the English speakers. The Spanish versions of these tests on the Batería III Woodcock–Muñoz (Muñoz-Sandoval, Woodcock, McGrew, & Mather, 2005) were administered to the Spanish speakers. Participants also completed language background and demographics questionnaires.

In session 2, participants completed the focal experimental tasks. There were two study–test cycles, one with study under full attention and one with study under a cognitive load. In both conditions, participants were instructed to read each word aloud and try to commit it to memory. Words appeared on the screen one at a time for 2 s each, with 1 s between words. The sequence contained a block of 40 high-frequency words and a block of 40 low-frequency words in counterbalanced order (as in Diana & Reder, 2006). The blocking procedure was used to be sure that frequency effects were due

to frequency per se and not due to differential allocation of attention to high- and low-frequency items. Also, this made it more like the pure-list condition used for the language manipulation in Experiment 2. Four medium-frequency filler words were included at the beginning, middle, and end of each sequence. Under cognitive load conditions, participants performed concurrently an auditory 2-back task with single digits from 1 to 5 as stimuli. One digit was presented during each gap between word presentations, and participants pressed *yes* or *no* buttons to indicate whether each digit matched the digit two before it. A yes–no recognition test immediately followed each study sequence. At test, the 80 studied words were randomly intermixed with 40 high- and 40 low-frequency foil words, thus making 160 test trials.

Results

Hit rates

Hit rates (see Table 2) were submitted to a 2 (language group) $\times$ 2 (cognitive load condition) $\times$ 2 (frequency) mixed ANOVA. Hit rates were lower under cognitive load conditions ($M = .56$) than under full-attention conditions ($M = .68$), $F(1, 62) = 52.62$, $MSE = .0197$, $p < .001$. Hit rates were not reliably different for low-frequency words ($M = .61$) and high-frequency words ($M = .63$), $F(1, 62) = 2.86$, $MSE = .0131$, $p = .096$. This can be explained in part by a significant interaction of cognitive load and frequency effects, $F(1, 62) = 11.22$, $MSE = .0079$, $p = .001$. Specifically, under full attention, hit rates for high-frequency ($M = .68$) and low-frequency ($M = .69$) words did

Table 2 Mean (*SE*) recognition performance in Experiment 1

Group	Attention condition	Word frequency	Hit rate (%)	FA rate (%)	<i>d'</i>
English monolinguals	Full attention	High frequency	73 (3)	29 (4)	1.34 (0.12)
		Low frequency	78 (3)	13 (2)	2.15 (0.14)
	Cognitive load	High frequency	60 (3)	45 (3)	0.44 (0.06)
		Low frequency	56 (3)	30 (3)	0.82 (0.11)
Spanish monolinguals	Full attention	High frequency	63 (3)	33 (3)	0.87 (0.09)
		Low frequency	60 (3)	18 (2)	1.32 (0.12)
	Cognitive load	High frequency	58 (3)	50 (3)	0.20 (0.06)
		Low frequency	50 (2)	35 (2)	0.41 (0.06)

not differ, $F < 1$, but under a cognitive load, hit rates were higher for high-frequency words ($M = .59$) than for low-frequency words ($M = .53$), $F(1, 62) = 14.052$, $MSE = .009$, $p < .001$. The English speakers ($M = .67$) had higher hit rates than the Spanish speakers ($M = .58$), $F(1, 62) = 9.72$, $MSE = .0560$, $p = .003$. There was a significant interaction of language group and cognitive load condition, $F(1, 62) = 15.59$, $MSE = .0218$, $p < .001$, such that the effect of cognitive load condition ($M = .58$ for cognitive load vs. $M = .76$ for full attention) was stronger for the English speakers, $F(1, 31) = 45.836$, $MSE = .021$, $p < .001$, than for the Spanish speakers ($M = .54$ for cognitive load vs. $M = .62$ for full attention), $F(1, 31) = 11.297$, $MSE = .018$, $p = .002$. No other interaction involving language group reached statistical significance, $ps > .05$.

False alarm rates

False alarm rates (see Table 2) were submitted to a 2 (language group) $\times$ 2 (cognitive load condition) $\times$ 2 (frequency) mixed ANOVA. False alarm rates were higher under cognitive load conditions ($M = .40$) than under full-attention conditions ($M = .23$), $F(1, 62) = 94.27$, $MSE = .0187$, $p < .001$. False alarm rates were lower for low-frequency words ($M = .24$) than for high-frequency words ($M = .39$), $F(1, 62) = 126.73$, $MSE = .0119$, $p < .001$. Cognitive load and frequency effects on false alarm rates did not interact, $F < 1$. Language group had no reliable effect, $F(1, 62) = 2.69$, $MSE = .0544$, $p = .106$, nor did it interact with the effects of cognitive load or frequency on false alarm rates, $Fs < 1$.

Discrimination

Detection statistic d' was calculated in each condition for each participant (see Table 2, Fig. 1), and these values were submitted to a 2 (language group) $\times$ 2 (cognitive load condition) $\times$ 2 (frequency) mixed ANOVA. Performing a concurrent task at study impaired recognition performance, in that discrimination was worse under a cognitive load ($M = .47$) than under full attention ($M = 1.42$), $F(1, 62) = 152.61$, $MSE = .381$, $p < .001$. Discrimination was better for low-frequency ($M = 1.18$) than for

high-frequency ($M = .71$) words, $F(1, 62) = 99.98$, $MSE = .137$, $p < .001$. These effects interacted, $F(1, 62) = 20.78$, $MSE = .091$, $p < .001$. Specifically, the low-frequency advantage was stronger under full attention ($M = 1.74$ for LF vs. $M = 1.11$ for HF), $F(1, 62) = 92.607$, $MSE = .139$, $p < .001$, than under a cognitive load ($M = .62$ for LF vs. $M = .32$ for HF), $F(1, 62) = 30.587$, $MSE = .089$, $p < .001$. English speakers ($M = 1.19$) exhibited higher discrimination scores than did Spanish speakers ($M = .70$), $F(1, 62) = 23.75$, $MSE = .648$, $p < .001$. Language group interacted with cognitive load condition $F(1, 62) = 4.47$, $MSE = .381$, $p = .039$, in that the effect of cognitive load condition was stronger for English speakers ($M = .63$ for cognitive load vs. $M = 1.75$ for full attention), $F(1, 32) = 87.819$, $MSE = .454$, $p < .001$, than for Spanish speakers ($M = .31$ for cognitive load vs. $M = 1.10$ for full attention), $F(1, 32) = 64.856$, $MSE = .308$, $p < .001$. An interaction of language group and word frequency, $F(1, 62) = 8.09$, $MSE = .137$, $p = .006$, indicated that the effect of word frequency was stronger for English speakers ($M = 1.49$ for low frequency vs. $M = .89$ for high frequency), $F(1, 31) = 66.409$, $MSE = .170$, $p < .001$, than for Spanish speakers ($M = .87$ for low frequency vs. $M = .54$ for high frequency),

Fig. 1 Recognition performance in Experiment 1 as a function of language group, attention condition, and word frequency

$F(1, 31) = 33.761$, $MSE = .104$, $p < .001$. The three-way interaction did not approach significance, $F(1, 62) = 1.66$, $MSE = .091$, $p = .202$.

Discussion

Under conditions of full attention, English speakers had higher hit rates and lower false alarm rates for low-frequency words than for high-frequency words, thus exhibiting the mirror effect, as in previous research. Although Spanish speakers did not show the hit rate advantage for low-frequency words, they did show the expected low-frequency advantage in discrimination scores. Imposing a cognitive load at study impaired recognition performance, with lower hit rates and higher false alarm rates than under full-attention conditions, consistent with previous research. The frequency effect was diminished under conditions of cognitive load, in that low-frequency words were hurt more than high-frequency words by the cognitive load, as in Diana and Reder (2006). The mirror effect was eliminated under conditions of cognitive load, in that both hits and false alarms were lower for low-frequency than for high-frequency words. However, the low-frequency advantage in discrimination held up even under a cognitive load.

Experiment 2

Experiment 2 was conducted to determine whether the cognitive load or word frequency conceptualization of L2 memory would better explain recognition memory performance in bilinguals. In Experiment 2, English-dominant and Spanish-dominant Spanish–English bilinguals completed recognition memory tests in both of their languages under full-attention conditions and intentional encoding. Study sequences contained both high- and low-frequency words and were identical to the sequences used for the monolingual participants in Experiment 1. Cognitive load and word frequency approaches to L2 processing lead to contrasting predictions for bilingual performance, as evidenced in Experiment 1. The cognitive load effect suggests that

bilingual recognition performance will be worse in L2 than in L1. In contrast, the frequency effect suggests that recognition performance will be better in L2 than in L1.

Method

Participants

Participants were 32 English-dominant and 32 Spanish-dominant bilinguals. The English-dominant participants and some of the Spanish-dominant participants were students at UTEP; the other Spanish-dominant participants were students at UACJ. Characteristics of the participants are given in Table 1. Bilingual status and language dominance were verified using the short forms of the WMLS–R in English and Spanish.

Design, materials, and procedure

The experimental conditions formed a 2 (language group) $\times$ 2 (language) $\times$ 2 (frequency) mixed design. Half of the participants were English-dominant and half were Spanish-dominant bilinguals. Language (English or Spanish) and word frequency (high or low) were manipulated within subjects. The materials were the same as in Experiment 1. The procedure was modified by having participants perform the study–test cycle once in English and once in Spanish (in counterbalanced order), both times under full-attention conditions. Thus, 80 words were studied in each language, 40 high frequency and 40 low frequency.

Results

Hit rates

Hit rates (see Table 3) were submitted to a 2 (dominance group) $\times$ 2 (language) $\times$ 2 (frequency) mixed ANOVA. Here, language was recoded to reflect the L1 and L2 of each participant. Hit rates were higher for L2 words ($M = .74$) than for L1 words ($M = .70$), $F(1, 62) = 5.90$, $MSE = .0151$, $p = .018$. Hit rates were higher for low-frequency words ($M = .76$) than for high-frequency words ($M = .68$), $F(1, 62) = 32.84$,

Table 3 Mean (*SE*) recognition performance in Experiment 2

Group	Language	Word frequency	Hit rate (%)	FA rate (%)	<i>d'</i>
English dominant	English	High frequency	67 (3)	32 (4)	1.07 (0.13)
		Low frequency	72 (3)	16 (3)	1.87 (0.16)
	Spanish	High frequency	69 (3)	28 (3)	1.22 (0.11)
		Low frequency	77 (3)	14 (2)	2.11 (0.16)
Spanish dominant	English	High frequency	69 (3)	24 (2)	1.35 (0.12)
		Low frequency	81 (2)	11 (1)	2.33 (0.13)
	Spanish	High frequency	68 (3)	28 (2)	1.15 (0.09)
		Low frequency	75 (3)	17 (2)	1.93 (0.15)

$MSE = .132$, $p < .001$. The interaction of language and frequency effects approached significance, $F(1, 62) = 3.60$, $MSE = .0081$, $p = .062$, suggesting a stronger frequency effect in L2 ($M = .79$ for low frequency vs. $M = .69$ for high frequency) than in L1 ($M = .74$ for low frequency vs. $M = .68$ for high frequency). Dominance group had no effect, nor did it interact with the effects of language or frequency, $F_s < 1$.

False alarm rates

False alarm rates (see Table 3) were submitted to a 2 (dominance group) $\times$ 2 (language) $\times$ 2 (frequency) mixed ANOVA. False alarm rates were lower for L2 ($M = .19$) than for L1 ($M = .23$), $F(1, 62) = 7.32$, $MSE = .0122$, $p = .009$. False alarm rates were lower for low-frequency words ($M = .14$) than for high-frequency words ($M = .28$), $F(1, 62) = 118.37$, $MSE = .0098$, $p < .001$. Language and frequency effects on false alarm rates did not interact, $F < 1$. Dominance group had no effect, $F < 1$, nor did it interact with the effects of language or frequency on false alarm rates, $p_s > .200$.

Discrimination

Values of d' (see Table 3, Fig. 2) were submitted to a 2 (dominance group) $\times$ 2 (language) $\times$ 2 (frequency) mixed ANOVA. Discrimination was better in L2 ($M = 1.76$) than in L1 ($M = 1.51$), $F(1, 62) = 9.18$, $MSE = .438$, $p = .004$. Discrimination was better for low-frequency ($M = 2.06$) than for high-frequency ($M = 1.20$) words, $F(1, 62) = 157.52$, $MSE = .303$, $p < .001$. These effects did not interact, $F(1, 62) = 2.05$, $MSE = .159$, $p = .157$. Dominance group did not have an effect on d' , nor did it interact with the effects of language or frequency, $F_s < 1$.

Fig. 2 Recognition performance in Experiment 2 as a function of language group, task language, and word frequency

Discussion

The mirror effect for high- and low-frequency words was observed, and it was evident in both L1 and L2. The primary new findings in this experiment were that following intentional encoding, L2 hit rates were higher than L1 hit rates, and L2 false alarm rates were lower than L1 false alarm rates. Thus, discrimination was greater for L2 words, and there was a language-based mirror effect.

Comparison of bilingual and monolingual performance

Bilingual and monolingual participants were comparable on age, parent education, visual-spatial reasoning ability, and fluid reasoning ability, as is shown in Table 1. Bilingual and monolingual performance (under full-attention conditions) were compared in two 2 (language group) $\times$ 2 (dominant language) $\times$ 2 (word frequency) mixed ANOVAs on discrimination scores. The first ANOVA compared bilingual L1 with monolingual performance. This analysis revealed a significant interaction of group and dominant language, $F(1, 124) = 9.838$, $MSE = .853$, $p = .002$, such that monolingual English speakers scored higher than monolingual Spanish speakers, $F(1, 62) = 17.717$, $MSE = .771$, $p < .001$, but bilingual English-dominant and Spanish-dominant speakers did not differ in their L1 performance, $F < 1$. No other effects or interactions involving group were reliable ($p_s > .10$).

The second ANOVA compared bilingual L2 with monolingual performance and showed a significant effect of group, with bilinguals outperforming monolinguals, $F(1, 124) = 8.674$, $MSE = .833$, $p = .004$. The interaction of group and frequency, $F(1, 124) = 7.300$, $MSE = .199$, $p = .008$, showed a larger low-frequency advantage for bilinguals in L2, $F(1, 62) = 108.087$, $MSE = .259$, $p < .001$, than for monolinguals, $F(1, 62) = 92.607$, $MSE = .139$, $p < .001$. The interaction of group and language dominance, $F(1, 124) = 13.272$, $MSE = .833$, $p < .001$, showed that while English-speaking monolinguals scored higher than Spanish-speaking monolinguals, $F(1, 62) = 15.575$, $MSE = .771$, $p < .001$, English-dominant and Spanish-dominant bilinguals did not differ in their L2 performance, $F(1, 62) = 1.131$, $MSE = .895$, $p = .292$. An unexpected three-way interaction of group, language dominance, and word frequency was reliable but not central to the hypotheses of the study, $F(1, 124) = 4.113$, $MSE = .199$, $p = .045$.

General discussion

There were two central new findings. First, following intentional encoding, bilingual recognition performance was more accurate in L2 than in L1, with both higher hit rates and lower false alarm rates in L2. Second, bilingual recognition

performance in L2 was more accurate than monolingual performance. Both effects were shown using lists that were pure with respect to language; therefore, the effects cannot be due to preferential processing of L2 items. The advantage for bilingual L2 performance over bilingual L1 and monolingual performance is consistent with previous research showing a recognition advantage for low-frequency over high-frequency words (e.g., Balota & Neely, 1980; MacLeod & Kampe, 1996).

This performance pattern can be explained by an adaptation of Reder's source-of-activation-confusion theory (Buchler & Reder, 2007; Diana & Reder, 2006). On the basis of this theory, we conclude that the mirror effect for bilingual L1 and L2 recognition memory arises for two reasons. First, because L2 words are preexperimentally associated with fewer episodic contexts, there is less contextual interference when L2 words are recollected, leading to higher hit rates. Second, L2 words have a lower baseline familiarity level, which makes them less vulnerable to familiarity-based false alarms. Thus, both the fan factor and the base factor of the source-of-activation-confusion theory apply to the domain of bilingual language dominance.

The *weaker-links* hypothesis or *frequency-lag* hypothesis suggests a common mechanism for word frequency and language proficiency effects for lexical processing in monolinguals and bilinguals (Gollan et al., 2008). The idea is that a bilingual has fewer lifetime exposures to L2 words than to the corresponding L1 words, which in turn are fewer than the number of exposures that a monolingual speaker would have. Therefore, L2 words are functionally of lower frequency than L1 words, a notion that is supported by the bilingual L2 advantage over bilingual and monolingual L1 performance in recognition. By the same logic, we also expected an advantage for bilingual L1 performance over monolingual performance, but any such difference was too small to detect. Consistent with findings for other tasks that rely primarily on comprehension rather than production (Gollan et al., 2011), we did not detect an interaction of language and frequency in recognition performance.

The central findings of this study do not rule out the influence of cognitive resource limitations on L2 memory processing. First of all, we cannot be sure that performing a concurrent working memory task (as in Experiment 1) and thinking in a less fluent language cause the same type of cognitive resource limitation. Second, research on basic memory processes suggests that although recall is particularly dependent on relational processing (Hunt & Einstein, 1981) or integration of items with each other and with information in long-term memory (Mandler, 1980), recognition is based in large part on specific-item processing (Hunt & Einstein, 1981) or familiarity based on within-item integration (Mandler, 1980). Divided attention appears to interrupt associative processing and, therefore, has larger effects on 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). Thus, the greater cognitive load associated with processing L2 may indeed produce a deficit in forming interitem associations, but such a deficit would not be expected to have a large impact on recognition memory.

Like low-frequency words, L2 words have an advantage in recognition because of their episodic distinctiveness. The results of the present study indicate that L2 performance is well characterized by the word frequency conceptualization of L2 memory in the domain of recognition memory. Further research will be needed to determine whether this conceptualization applies equally well to other bilingual memory phenomena.

Author Note This research was supported by NSF grant BCS-0924905 to the first author. We thank Gabriela Durán, Betsabee Lara, and Marlene González for assistance in data collection. Preliminary results were presented at the 51st Annual Meeting of the Psychonomic Society.

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.
- 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.
- Buchler, N. E. G., & Reder, L. M. (2007). Modeling age-related memory deficits: A two-parameter solution. *Psychology and Aging, 22*, 104–121.
- 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.
- Durgunoglu, A. Y., & Roediger, H. L. (1987). Test differences in accessing bilingual memory. *Journal of Memory and Language, 26*, 377–391.
- Francis, W. S., & Gutiérrez, M. (2012). Bilingual recognition memory: Stronger performance but weaker levels-of-processing effects in the less fluent language. *Memory & Cognition, 40*, 496–503.
- 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.
- Gollan, T. H., Slattery, T. J., Goldenberg, D., Van Assche, E., Duyck, W., & Rayner, K. (2011). Frequency drives lexical access in

- reading but not in speaking: The frequency-lag hypothesis. *Journal of Experimental Psychology: General*, 140, 186–209.
- 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.
- 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.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- 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 and Cognition*, 10, 33–42.
- Muñoz-Sandoval, A. F., Woodcock, R. W., McGrew, K. S., & Mather, N. (2005). *Bateria III Woodcock-Muñoz*. Itasca, IL: Riverside Publishing.
- 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: Human Experimental Psychology*, 53A, 609–625.
- 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.
- 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.
- Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson III*. Itasca, IL: Riverside Publishing.
- Woodcock, R. W., Muñoz-Sandoval, A. F., Ruef, M. L., & Alvarado, C. G. (2005). *Woodcock-Muñoz Language Survey-Revised*. Rolling Meadows, IL: Riverside Publishing.