



## Research Article

**Cite this article:** Francis, W.S., & Nájera, O.I. (2024). Bilingual proficiency effects in paired-associate learning of vocabulary in an unfamiliar language. *Bilingualism: Language and Cognition*, 1–18. <https://doi.org/10.1017/S1366728924000130>

Received: 25 May 2023  
Revised: 30 January 2024  
Accepted: 4 February 2024

**Keywords:**  
vocabulary learning; bilingualism; language proficiency; paired-associate learning

**Corresponding author:**  
Wendy Francis;  
Email: [wfrancis@utep.edu](mailto:wfrancis@utep.edu)

### Abstract

We investigated three aspects of paired associate learning of vocabulary in an unfamiliar language: monolingual-bilingual differences, effects of dominance and language proficiency, and the possible role of associative strategies. Spanish–English bilinguals (48 English-dominant and 48 Spanish-dominant) and English-speaking monolinguals ( $n = 48$ ) learned Swahili–English and Swahili–Spanish word pairs. Learning was assessed using cued recall (Swahili cue or Swahili response) and associative recognition tests. English-dominant bilinguals did not outperform English monolinguals on any learning measure. Cued recall accuracy was higher when learning through the dominant language than through the non-dominant language, whether the Swahili words were cues or responses. Proficiency scores in the known language were positively correlated with cued recall accuracy, whether the cue or the response was in Swahili, indicating that proficiency effects occurred not in retrievability of known words but in learning of associations. Bilingual and monolingual participants did not differ in their reported use of associative strategies.

### Introduction

Vocabulary in an unfamiliar language can be learned using paired associations of the new unfamiliar words with their translations in a known language. For example, in paired-associate learning of Swahili vocabulary, an English speaker would study an unfamiliar word in Swahili (*mbwa*) presented together with its translation equivalent in English (*dog*) and attempt to commit the pairings to memory. Learning can be tested using cued recall or associative-recognition tests. In a CUED-RECALL test, one word in the pair is presented and the other must be generated. Typically, the cued-recall test requires RECEPTIVE RETRIEVAL, in that the unfamiliar word is presented as a cue to recall the familiar word. For example, *mbwa* could be presented and the expected response would be *dog*. The cued-recall task can be made more difficult by requiring PRODUCTIVE RETRIEVAL, in which the familiar word as a cue to recall the unfamiliar word (e.g., de Groot & Keijzer, 2000; Griffin & Harley, 1996; Schneider et al., 2002). For example, *dog* would be presented, and the expected response would be *mbwa*. In an associative-recognition test, pairs of words are presented for verification if it is a correct pairing or rejection if it is an incorrect pairing. For example, if *mbwa-dog* were presented, the response would be *yes*; if *mbwa-table* were presented, the response would be *no*.

In the present study, we investigated three aspects of paired associate learning of unfamiliar vocabulary: (1) monolingual-bilingual differences in learning; (2) associations of language dominance and proficiency with learning; and (3) the possible role of learning strategies. Spanish–English bilinguals learned Swahili words through both of their known languages, and learning was measured using cued recall tests in both directions and associative recognition tests.

### Bilingual-monolingual comparisons

Several studies have reported that cued recall and associative recognition of foreign vocabulary were more accurate for bilinguals using their more proficient language than for monolinguals (Bogulski et al., 2018; Kaushanskaya, 2012; Kaushanskaya & Marian, 2009; Kaushanskaya & Rehtzigel, 2012; Van Hell & Mahn, 1997). One study similarly showed more accurate performance in multilinguals than in bilinguals (Papagno & Vallar, 1995). Several explanations for these effects have been offered (see Tsuboi & Francis, 2020 for a review). Among these are that bilinguals may have greater phonological working memory or phonological long-term knowledge (Kaushanskaya, 2012; Papagno & Vallar, 1995; Van Hell & Mahn, 1997), enhanced selective attention or executive function (Kaushanskaya, 2012; Kaushanskaya & Marian, 2009), more effective encoding strategies (Kaushanskaya & Rehtzigel, 2012), and greater experience

© The Author(s), 2024. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.



This article has earned badges for transparent research practices: Open Data and Open Materials. For details see the Data Availability Statement.

in associating novel labels with concepts (Kaushanskaya, 2012). Other possible reasons to expect bilinguals to learn vocabulary more efficiently include more efficient association of previously meaningless information to familiar meaningful information and a greater vocabulary to which associations can be made (Tsuboi & Francis, 2020).

Although group differences were large ( $\eta_p^2 \geq .15$ ) and consistent, five studies had one or more features that would bias performance in favor of the bilingual (or multilingual) participants (see Tsuboi & Francis, 2020 for a more detailed review). First, multilingual self-selected language majors were compared to bilinguals whose less proficient language was a standard academic requirement (Papagno & Vallar, 1995). Second, bilingual dominant-language speakers took 60% longer per item to study than monolinguals or nondominant-language speakers (Bogulski et al., 2018). Third, the bilinguals were first-language English speakers in the U.S. self-selected for K-12 education enriched with second-language classroom immersion and had English vocabulary or working memory spans greater than or equal to the monolingual participants (Kaushanskaya, 2012; Kaushanskaya & Marian, 2009; Kaushanskaya & Rehtzigel, 2012). Matching monolingual and bilingual groups on these factors results in non-representative bilingual samples with particularly high language-learning ability (Prior & MacWhinney, 2010). These five studies also had small bilingual samples ( $n \leq 25$ ). In the only study with adequate sample size, monolingual and bilingual groups learned different foreign languages which may not have had equal difficulty (Van Hell & Mahn, 1997). In this last study and some of the others, group matching on age, education, and/or SES was inadequate or not reported. None of these studies had the rigor necessary to draw definitive conclusions about monolingual-bilingual differences.

In contrast, we recently reported a larger and more rigorously controlled study in which monolingual, English-dominant, and Spanish-dominant young adults were matched on age, education, nonverbal cognitive ability, and SES (Tsuboi & Francis, 2020). With 48 participants per group, cued recall and associative recognition accuracy for Japanese-English word pairs showed no group differences. The use of associative strategies (e.g., mediator, sentence) to learn the words was reported more often by bilingual than by monolingual participants (74% vs. 46%), but participants who reported using associative strategies did not have more accurate performance.

As in previous comparisons of monolingual and bilingual cued recall, the Tsuboi and Francis (2020) study used a receptive retrieval task in which the unfamiliar words were used to cue recall of the familiar words. However, it is also informative to consider productive retrieval tasks in which familiar words are presented as cues to recall the corresponding unfamiliar words. Cued recall with productive retrieval is more difficult and depends to a greater degree on phonological memory (e.g., de Groot & Keijzer, 2000; Griffin & Harley, 1996; Schneider et al., 2002). Learning strategies may also be more important for productive retrieval. The greater dependency on phonological memory or the greater potential impact of associative strategy use might be expected to favor bilingual relative to monolingual learners. Published studies using this reversed direction for cued recall include no comparisons of monolingual and bilingual performance. There is one multilingual vs. bilingual study that used productive retrieval (Papagno & Vallar, 1995), but the findings were inconclusive because of self-selection bias and small sample sizes ( $ns = 10$ ). The present study includes both receptive and

productive cued-recall tasks to determine whether group differences emerge in the more difficult productive retrieval task.

### *Language dominance and proficiency*

Language dominance and language proficiency are related but distinct constructs. In general, language proficiency is skill in a particular language, whereas language dominance refers to the relative skill in two languages within an individual speaker. These factors were examined in only two of the reported bilingual studies of foreign vocabulary learning. First, dominant-language speakers had more accurate performance than nondominant-language speakers (Bogulski et al., 2018), but the comparison was compromised by the longer study times of the dominant-language speakers. In the same study, limited unbiased evidence was obtained, in that subjective proficiency ratings in a small sample of Spanish-English bilinguals correlated with accuracy. Specifically, bilinguals who were more English dominant had more accurate learning of Dutch-English paired associates than bilinguals who were more Spanish-dominant (Bogulski et al., 2018).

In a recent study with a substantially larger sample (Tsuboi & Francis, 2020), objectively assessed proficiency in the known language that was paired with the unfamiliar language correlated reliably with accuracy in cued recall and associative recognition. Specifically, in Spanish-English bilinguals learning Japanese vocabulary through Japanese-English translation pairs, objective English proficiency scores were positively correlated with accuracy on cued recall and associative recognition tests. Similarly, in Japanese-English bilinguals learning Spanish vocabulary through Japanese-Spanish translation pairs, objective Japanese proficiency scores were positively correlated with accuracy on cued recall and associative recognition tests. In both samples, associations of proficiency in the known language not used for the learning task (Spanish and English, respectively) were less conclusive, with some significant effects in the Spanish-English bilinguals but not in the Japanese-English bilinguals. These findings suggest that proficiency in the language paired with the unfamiliar language was an important factor in learning. However, learning of Japanese-English translation pairs did not differ between English-dominant and Spanish-dominant participants (Tsuboi & Francis, 2020), possibly because of the substantial overlap in English proficiency distributions between the two groups. In the present study, Spanish-English bilingual participants learned Swahili vocabulary through both English and Spanish.

In cued recall with receptive retrieval, generating responses in a less proficient language requires retrieval of words that are more difficult to access. The relative difficulty of access to the known language words would therefore be a plausible explanation of the proficiency effects observed in cued recall with receptive retrieval. Even the proficiency effects observed in associative recognition could be explained in this way, if we adopt a recall-to-reject account of associative recognition (Anderson, 1974; Clark, 1992; Mandler, 1980). In contrast, with productive retrieval, the words to be retrieved would be equally unfamiliar across levels of proficiency in the cue language. Therefore, if proficiency effects were to extend to this direction of cued recall, they could not be attributed to differences in difficulty accessing the unfamiliar response words. Because cued recall requires retrieval of both the association formed between the two words of the pair and retrieval of the appropriate response word (e.g., Madan et al., 2010), any such effects would have to be based on the

strength of the association formed between the two words in each pair. To distinguish between these explanations, the present study included cued recall tests with productive retrieval.

### Present study

The overarching goal of the present study was to better understand the effects of bilingualism and bilingual proficiency in an experiment that followed the rigor of the Tsuboi and Francis (2020) study but enhanced the likelihood that effects would be detected in two ways. First, the study included cued-recall tests in the more difficult direction, where monolingual-bilingual differences might be magnified and have not previously been compared. Second, each bilingual participant learned vocabulary through both of their languages. Specifically, English-dominant and Spanish-dominant Spanish-English bilinguals learned Swahili-English and Swahili-Spanish pairings in separate phases of the experiment (English-speaking monolinguals learned Swahili vocabulary through Swahili-English pairings only.)

The first goal was to rigorously compare monolingual and bilingual learning of vocabulary in an unfamiliar language and explore the plausibility of associative strategy use as a mechanism to explain any differences observed. Based on our previous result (Tsuboi & Francis, 2020), we did not expect to find differences between monolingual and bilingual accuracy on cued recall tests with receptive retrieval or on associative recognition tests. In the more difficult cued recall test with productive retrieval, we hypothesized that bilinguals would retrieve the unfamiliar words more efficiently due to a higher rate of associative strategy use or a larger vocabulary (e.g., Gollan et al., 2008) with which to make such associations. Therefore, we predicted that bilingual performance would be enhanced relative to monolingual performance when recalling Swahili words in response to English or Spanish cues.

To address the possible role of associative strategies in any group differences observed, we analyzed self-reported learning strategies. We assessed group differences in the reported use of associative strategies and whether the use of these strategies was associated with learning. We predicted that bilinguals would report using more associative strategies than monolinguals as in our previous research (Tsuboi & Francis, 2020). Although the reported use of these strategies in that study was not associated with cued recall performance with receptive retrieval, we hypothesized that associative strategies would be helpful for productive retrieval. We therefore predicted that participants who reported using associative strategies would have greater accuracy in the difficult direction of cued recall than participants who did not.

The second goal was to clarify the relationships of language dominance and proficiency to learning, as measured by accuracy in cued recall and associative recognition tests. For each participant, the dominant language was considered to be the language with the higher proficiency score. We hypothesized that learning would be more efficient when unfamiliar words were paired with the dominant language, because of having a larger vocabulary in that language with which to associate the novel Swahili words. Specifically, we predicted that cued recall with receptive retrieval and associative recognition would be more accurate when Swahili words were paired with their translations in the dominant language. Similarly, we hypothesized that higher proficiency in the language through which unfamiliar vocabulary was learned would be associated with greater learning, also because of having a larger vocabulary with which to associate the novel Swahili

words. We also hypothesized that proficiency in other languages would not be associated with learning. Specifically, for cued recall with receptive retrieval and associative recognition, we predicted that English performance would be positively correlated with English proficiency but not Spanish proficiency. Similarly, we predicted that Spanish performance would correlate with Spanish but not English proficiency. We also hypothesized that the effects of dominance and proficiency do not depend on the relative retrievability of familiar response words. Instead, bilinguals with higher proficiency in the task language would form stronger associations at encoding that would be easier to retrieve at test. Thus, in the more difficult cued recall test with productive retrieval, if the effects of dominance or proficiency are in the encoding of new associations, we would expect the same pattern of effects as for receptive retrieval. In contrast, if proficiency effects are based on the retrievability of known words, then no effects of dominance or proficiency would be expected in productive retrieval.

## Method

### Power and sample size

Reported monolingual-bilingual differences in previous studies were very large effects (cued recall: Mean  $d = 1.11$ ; associative recognition: Mean  $d = 1.05$ ; averaged across effect sizes cited in Table 1 of Tsuboi and Francis (2020), excluding that study and comparisons in which bilinguals learned through their nondominant language). Even the smallest reported difference was a medium-to-large effect ( $d = .63$ ; Kaushanskaya & Reetzgel, 2012). Nevertheless, we wanted to have 80% power to detect medium sized pairwise group differences ( $d = .50$ ) tested as single-df comparisons in a three-group design. A power analysis in G\*Power showed that this would require 128 participants in total, at least 43 participants for each of the three language groups (note that because power analyses for logistic mixed-effects regression are not well developed, we used a power analysis based on ANOVA.) Due to counterbalancing considerations requiring a multiple of 8 in each group, we tested 48 participants in each language group, which yields power of 85% to detect a medium group difference.

We also anticipated examining the effects of language dominance and proficiency in a continuous manner. In previous studies, the correlations of dominance scores and accuracy were medium to large. In the most similar study (Tsuboi & Francis, 2020, Experiment 1), correlations of oral language scores with cued recall and associative recognition scores in Spanish-English bilinguals ranged from .311 to .447. With 96 bilingual participants, the present experiment has 85% power to detect a medium correlation ( $r = .30$ ).

### Participants

The participants were English-speaking monolinguals, English-dominant Spanish-English bilinguals, and Spanish-dominant Spanish-English bilinguals ( $N = 144$ ), with 48 participants in each language group. All participants were recruited at the University of Texas at El Paso (UTEP), and they received course credits or \$10 per hour for research participation. Seven participants with complete experimental and assessment data were excluded: 1 because of low English proficiency not detected until after the experiment was complete, 1 because of extremely low cognitive scores, and 5 because they were inadvertently run

**Table 1** Participant Characteristics

Characteristic	English Monolinguals <sup>a</sup> ( <i>n</i> = 48)	English-Dominant Bilinguals ( <i>n</i> = 48)	Spanish-Dominant Bilinguals ( <i>n</i> = 48)
Median Age	20.0 (6.7)	20.0 (3.5)	20.5 (6.0)
First Language Exposed <sup>b</sup>	Spanish 4% English 83% Both 6%	Spanish 67% English 6% Both 27%	Spanish 96% English 2% Both 2%
Median Age of English Acquisition	1.0 (1.4)	5.0 (2.1)	7.0 (5.5)
Frequency of English Usage			
Every day or several days per week	94%	100%	96%
Weekly	0%	0%	2%
Every two weeks or less	0%	0%	2%
Frequency of Spanish Usage			
Every day or several days per week	21%	96%	98%
Weekly	21%	2%	2%
Every two weeks or less	52% <sup>c</sup>	2%	0%
WMLS-R Language Assessments <sup>d</sup>			
English Picture Vocabulary	17.1 (5.3)	15.4 (4.7)	10.2 (2.6)
English Verbal Analogies	22.3 (7.8)	20.1 (8.7)	14.8 (7.0)
English Oral Language Composite	19.0 (6.4)	16.5 (5.8)	11.3 (2.6)
Spanish Picture Vocabulary	2.9 (1.3)	10.1 (2.3)	13.9 (3.6)
Spanish Verbal Analogies	5.0 (1.4)	14.1 (7.7)	22.6 (8.8)
Spanish Oral Language Composite	3.2 (1.7)	11.1 (3.8)	16.2 (5.6)
WJ-III Cognitive Assessments <sup>e</sup>			
Visual-Spatial Thinking Composite	14.3 (6.6)	16.4 (6.6)	16.0 (6.3)
Women	60%	60%	63%
Reported Latino Ethnicity	69%	100%	98%
Median Education	Some College	Some College	Some College
Median Highest Parental Education	Graduated College	Graduated College	Graduated College

Note: Values indicate means or percentages unless otherwise indicated; numbers in parentheses are standard deviations.

<sup>a</sup>Some percentages do not add to 100%, because several questionnaire responses were missing for three participants.

<sup>b</sup>Both indicates that reported ages of acquisition for the two languages were one year apart or less, indicating simultaneous acquisition

<sup>c</sup>2% biweekly, 6% monthly, 10% every few months, 15% once or twice per year, 19% less than once or twice per year

<sup>d</sup>Mean age-equivalency scores for performance on the WMLS-R (Woodcock et al., 2005). Age equivalency is the normative age in years at which performance would be average for a monolingual speaker.

<sup>e</sup>Scores indicate mean age-equivalency levels for performance on the WJ-III (Woodcock et al., 2001).

on the wrong form. All excluded participants were replaced to preserve counterbalancing. Participant self-reported characteristics and objective assessment scores are summarized in Table 1.

The objective assessments of proficiency in English and Spanish were the Picture Vocabulary and Verbal Analogies subtests of the Woodcock-Munoz Language Survey-Revised (WMLS-R; Woodcock et al., 2005). This instrument is an objective and standardized assessment normed with monolingual English and monolingual Spanish speakers in the Americas. Scores for each subtest were entered into the scoring program, which uses the scores to compute oral language composite scores for each language, which were then converted to age-equivalency scores. The age-equivalency score represents the average age at which a monolingual speaker achieves the same performance level and is reported in Table 1. However, for inferential analyses, calibrated scores (*W*) provided by the WMLS-R scoring program

were used, because they are closer to normally distributed (see Schrank et al., 2010 for technical information). It is important to note that these two subtests rely primarily on knowledge of individual words in the language and not grammar or connected language, which limits its scope as an index of proficiency.

The bilingual participants were students who had extensive experience using English and Spanish for substantial communication and had little to no difficulty conversing fluently with the experimenter in both languages. Nearly all of the students (96%) reported either learning Spanish first or learning English and Spanish simultaneously from early childhood. Current regular usage is evidenced by 98% reporting that they used both English and Spanish either every day or several days per week. Bilingual participants had to score at least 8 years on the Oral Language composite in both English and Spanish or demonstrate clear spoken proficiency in conversing with the experimenter with

a score of at least 7 years. Bilingual participants whose English score was higher than their Spanish score were classified as English-dominant bilinguals; those whose Spanish score was higher than their English score were classified as Spanish-dominant bilinguals. Their age-equivalency scores on the Oral Language composite averaged 16.3 years in their more proficient language and 11.2 in their less proficient language. Based on calibrated W scores, English-dominant bilinguals scored higher in English than the Spanish-dominant bilinguals,  $t(94) = 6.500$ ,  $p < .001$ , and Spanish-dominant bilinguals scored higher in Spanish than the English-dominant bilinguals,  $t(94) = 6.859$ ,  $p < .001$ . Also, the degree of dominance did not differ reliably across groups,  $t(94) = 1.080$ ,  $p = .283$ .

The monolingual group consisted of students who were FUNCTIONALLY MONOLINGUAL, in that they did not have fluent conversational skills in Spanish and had relatively brief and infrequent use of the little Spanish they knew; they were not able to converse fluently with the experimenters. The vast majority (83%) learned English as their only first language, and 12% were exposed to both languages early in life but did not become proficient in Spanish. English monolingual participants had to obtain an age-equivalency score of at least 8 years on the Oral Language composite in English, have no reported proficiency in any other languages, not be able to engage in Spanish conversation with the experimenter, and obtain scores less than 8 years in Spanish to qualify to participate in the present experiment. Their age-equivalency scores on the Oral Language composite averaged 19.0 in English (significantly higher than the English-dominant group,  $t(94) = 2.29$ ,  $p = .024$ ) and 3.2 in Spanish (significantly lower than the English-dominant group,  $t(94) = 17.01$ ,  $p < .001$ ).

As shown in Table 1, the three language groups were well matched on age, education, and socio-economic status, as indicated by the highest parent education level (Galobardes et al., 2006). Age differences among groups were not significant,  $F(2, 140) = .488$ ,  $MSE = 30.74$ ,  $p = .615$ ,  $\eta_p^2 = .007$ . Education for participants and parents was measured on a six-point ordinal scale (less than 8th grade, some high school, graduated high school, some college, graduated college, advanced degree). All participants were in the “some college” category. Reported parent education levels had a median of “graduated college” for all groups. Two subtests from the Woodcock-Johnson III Test of Cognitive Abilities (WJ-III; Woodcock et al., 2001) were administered: Spatial Relations and Picture Recognition. The scoring program used these subtests to compute a composite visuo-spatial reasoning score and convert all scores to age-equivalency scores. Here, although the English monolinguals scored somewhat lower numerically than the bilinguals, the difference was not statistically significant,  $F(1, 142) = 2.864$ ,  $p = .093$ ,  $\eta_p^2 = .02$ .

**Materials**

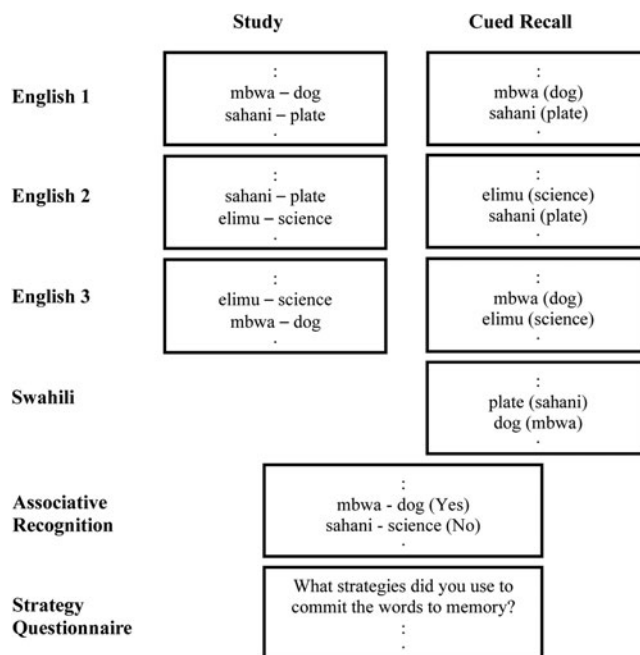
A set of 68 Swahili-English word pairs was selected from Nelson and Dunlosky (1994). The pairs were selected according to three criteria: 1) There was a clear Spanish translation equivalent; 2) The Swahili word did not resemble the English or Spanish translation equivalent; 3) The English and Spanish translation equivalents were deemed likely to be known in the less proficient language of a bilingual. The stimulus set included English, Spanish, and Swahili translation equivalents for 68 word concepts (see Appendix A). The mean letter length was 5.5 for the Swahili words, 5.6 for the English words, and 6.3 for the Spanish words.

These words were randomly divided into two sets matched on difficulty (based on norms from Nelson & Dunlosky, 1994) and on the proportion of concrete and abstract words (82% concrete, 18% abstract). The assignment of item sets to languages was counterbalanced across participants. Also, the half of each set that was assigned to be correctly or incorrectly paired in the associative recognition text was counterbalanced across participants. Incorrect pairings for this test were randomly generated.

**Procedure**

At the beginning of the session, following informed consent, the WMLS-R was administered in English and Spanish. In the main experiment, participants learned two different sets of 34 Swahili words. As illustrated in Figure 1, bilingual participants studied and were tested on Swahili-English translation pairs and Swahili-Spanish translation pairs in separate trial blocks, with the order of languages counterbalanced across participants. Monolingual participants studied and were tested on two sets of Swahili-English translation pairs in separate trial blocks. Between blocks, participants completed cognitive assessments and questionnaires.

Each language block began with three study-test cycles with Swahili cues and responses in the known language. A series of 34 Swahili-English (or Swahili-Spanish) word pairs were presented one at a time in random order for 8 sec each, during which participants were to say the pair aloud three times. The screen went blank for 1 sec between word pairs. Immediately after studying the last word pair, the cued recall task began. The 34 Swahili words were presented one at a time in random order, and the participant attempted to recall and say the



**Figure 1. Diagram of Experimental Procedure.** Participants completed three study-recall cycles, studying 34 Swahili-English word pairs and completing cued recall tests with Swahili cues for English responses. Next, they completed a cued recall test with English cues and Swahili responses and an associative recognition test with correct and incorrect Swahili-English pairings. Upon completion, they were given a strategy questionnaire. This procedure was repeated with Spanish-Swahili pairs (in counterbalanced order). Note that words in parentheses are expected cued recall responses.

corresponding English (Spanish) word aloud. The experimenter used a worksheet that contained the expected responses to mark correct and incorrect responses. Upon completion of the first cued recall test, the study phase for the next cycle began. The second and third study test cycles were the same as the first except that the items were presented in different random orders. Upon completion of the third study-test cycle, an additional cued recall test was given in which the 34 English (Spanish) words were presented one at a time in random order, and the participant was to recall the corresponding Swahili words.

Next, participants completed an associative recognition test. On this test, 34 word pairs were presented, 17 of the originally studied pairs and 17 incorrect pairings of studied Swahili and English (Spanish) words. Participants were to indicate whether each pairing was correct or incorrect by pressing a *yes* (*l*) or *no* (*z*) key on the computer keyboard. Upon completion of this final test, participants were given a questionnaire (adapted from Tsuboi & Francis, 2020) that asked them to explain the strategies that they used to commit the words to memory and later recall them.

The procedure as described was used for in-person administration of the protocol. About halfway through data collection, the protocol was adapted for remote administration during the pandemic, and 41 participants completed the remote version of the experiment in scheduled Zoom appointments with the experimenters. Participants were instructed to put Zoom into full-screen mode to avoid being distracted by other applications. The only change to the language assessments is that the picture vocabulary test items were displayed on the shared computer screen. Questionnaires were programmed and administered through QuestionPro. In the computerized experiment proper, PsyScope was used to present the stimuli on a shared screen. The only change was that on the associative recognition test, participants simply said yes or no aloud, and the experimenter entered their responses.

### Approach to analysis

The primary dependent variables were cued recall accuracy, associative recognition accuracy, and associative strategy use. The first set of analyses examined group differences in accuracy, with a focus on comparisons between monolinguals and bilinguals on learning of Swahili–English word pairs only. The second set of analyses, which include data only from the bilingual participants,

examined the effects of language dominance on accuracy using both Swahili–English and Swahili–Spanish word pairs. Here, the dominant language was operationalized as the language with the higher language proficiency score. The third set of analyses examined the associations of language proficiency scores with accuracy separately for bilingual and monolingual participants. Within each of these sets of analyses, logistic mixed-effects regression was used to analyze cued recall accuracy, and ANCOVAs on signal-detection measures were used for associative recognition accuracy. The final set of analyses focused on the use of associative strategies. Reported usage of the strategies was compared across groups using chi-square tests, and the associations of reported strategies with accuracy on the learning measures were analyzed using ANOVA.

## Results

### Accuracy

Mean accuracy scores are given in Table 2 for all groups, conditions, and measures. Cued recall data were analyzed using logistic mixed-effects regression in Jamovi (The jamovi project, 2022). We report the effects of interest in this section, and detailed results are provided in tables in Appendix B. In each analysis, we used the maximal random-effects structure that converged. Associative recognition data were analyzed using ANCOVAs on the signal detection measure  $d'$ . In LME and ANCOVA analyses comparing groups, we included cognitive scores as a covariate to remove the association of cognitive scores with accuracy from the error variance (the results of more traditional analyses using ANOVA are provided in the data repository for comparison with previous studies.)

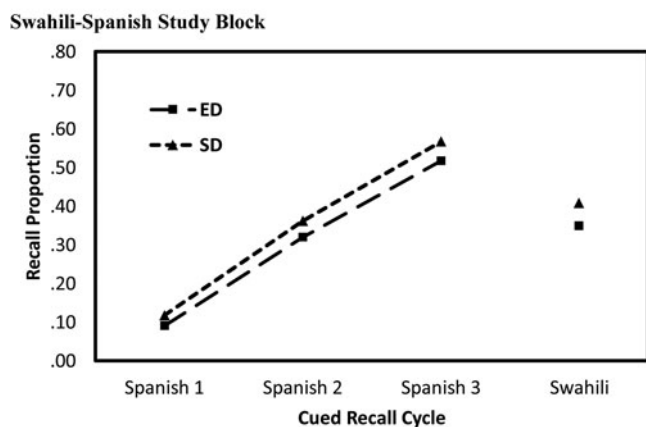
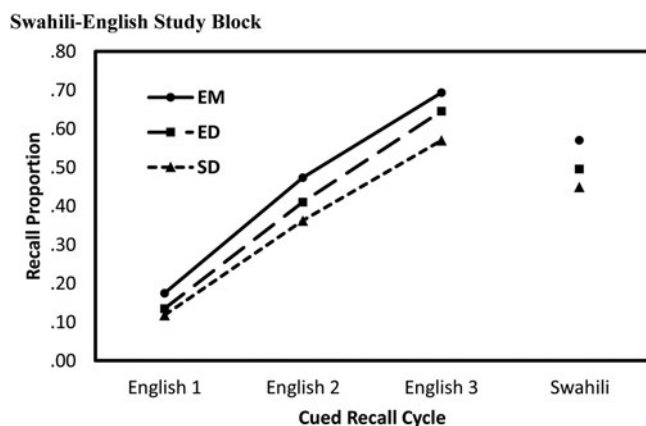
### Monolingual-bilingual comparisons

To compare the performance of the monolingual and bilingual groups, only data from English trial blocks were included in the analyses. Cued recall results are illustrated in Figure 2. For cued recall with receptive retrieval (Swahili cued, English responses), the fixed factors included language group, recall cycle (within-subjects), and their interaction. Cognitive assessment scores ( $W$  scores) were included as a covariate. The model included random intercepts for participants and items, random slopes across recall cycle for participants and items, and random slopes across

**Table 2** Mean (SE) Cued Recall and Associative Recognition Accuracy.

Language Group	Cued Recall			Associative Recognition				
	Task Language	Cycle 1 (proportion)	Cycle 2 (proportion)	Cycle 3 (proportion)	Swahili (proportion)	Hit rate (proportion)	FA rate (proportion)	$d'$
English Monolingual								
English	.175 (.019)	.473 (.037)	.693 (.035)	.570 (.037)	.909 (.016)	.125 (.018)	2.91 (.78)	
English Dominant								
English	.135 (.018)	.411 (.034)	.646 (.037)	.495 (.034)	.921 (.012)	.147 (.025)	2.97 (.16)	
Spanish	.090 (.014)	.320 (.033)	.517 (.040)	.349 (.030)	.890 (.016)	.185 (.026)	2.61 (.18)	
Spanish Dominant								
English	.118 (.016)	.363 (.030)	.570 (.037)	.449 (.033)	.901 (.018)	.156 (.019)	2.72 (.14)	
Spanish	.118 (.016)	.362 (.031)	.567 (.038)	.407 (.036)	.912 (.016)	.167 (.021)	2.78 (.17)	

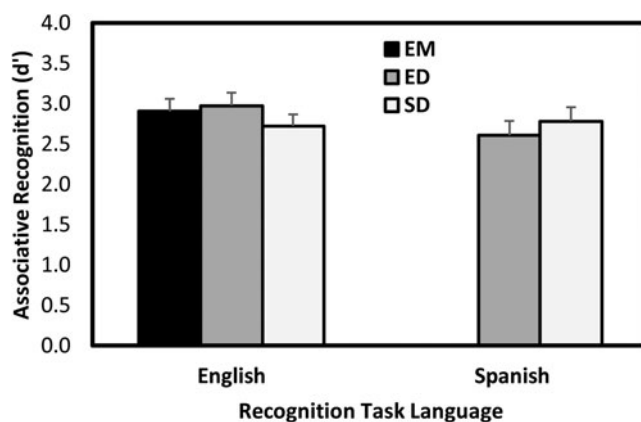
Note: Cued recall cycles 1, 2, and 3 refer to recall with Swahili cues; Swahili cued recall refers to recall with Swahili responses. FA = false alarm



**Figure 2. Cued Recall Accuracy.** Cued recall performance across three study-recall cycles with Swahili cues and one cycle with Swahili responses. EM = English monolingual, ED = English-dominant bilingual, SD = Spanish-dominant bilingual

language groups for items. Planned comparisons of groups showed that recall was more accurate for English monolinguals relative to Spanish-dominant bilinguals,  $b = .773$ ,  $SE = .277$ ,  $z = 2.796$ ,  $p = .005$ , but not relative to English-dominant bilinguals,  $b = .465$ ,  $SE = .273$ ,  $z = 1.705$ ,  $p = .088$ . Accuracy of the two bilingual groups did not differ ( $p > .2$ ). Accuracy improved across the three study-test cycles, as indicated by a significant linear trend,  $b = 2.408$ ,  $SE = .078$ ,  $z = 30.834$ ,  $p < .001$ . A significant interaction showed that the difference between monolingual and Spanish-dominant bilingual performance increased/decreased across the three cued recall cycles,  $b = .364$ ,  $SE = .152$ ,  $z = 2.391$ ,  $p = .017$ . Higher cognitive assessment scores were associated with higher accuracy,  $b = .021$ ,  $SE = .009$ ,  $z = 2.343$ ,  $p = .019$  (see Table B.1 for additional details.)

For cued recall with productive retrieval (English cues, Swahili responses), the fixed factor was language group. Cognitive assessment scores were included as a covariate. The model included random intercepts for participants and items and random slopes across language groups for items. Planned comparisons showed that English monolinguals had more accurate productive cued recall than English-dominant,  $b = .658$ ,  $SE = .310$ ,  $z = 2.125$ ,  $p = .034$ , and Spanish-dominant bilinguals,  $b = .801$ ,  $SE = .314$ ,  $z = 2.550$ ,  $p = .011$ . The accuracy of productive cued recall did not differ for the two bilingual groups ( $p > .5$ ). Cognitive assessment scores were not significantly associated with accuracy ( $p > .2$ ) (see Table B.2 for additional details.)



**Figure 3. Associative Recognition Accuracy.** Error bars indicate standard errors of the mean. EM = English monolingual, ED = English-dominant bilingual, SD = Spanish-dominant bilingual.

For the associative recognition data, hit rates and false alarm rates were obtained and used to compute the signal detection statistic  $d'$  (illustrated in Figure 3). The  $d'$  scores were submitted to a one-way between-subjects ANCOVA across language groups, with cognitive assessment scores as a covariate. No group differences were detected (English monolingual vs. English dominant:  $F(1, 140) = .01$ ,  $MSE = 1.097$ ,  $p = .928$ ,  $\eta_p^2 < .01$ ; English monolingual vs. Spanish dominant:  $F(1, 140) = .91$ ,  $MSE = 1.097$ ,  $p = .342$ ,  $\eta_p^2 < .01$ , English dominant vs. Spanish dominant:  $F(1, 140) = 1.56$ ,  $MSE = 1.097$ ,  $p = .213$ ,  $\eta_p^2 = .01$ ). Higher cognitive scores were associated with greater accuracy,  $F(1, 140) = 4.148$ ,  $MSE = 1.097$ ,  $p = .044$ ,  $\eta_p^2 = .029$ .

*Bilingual language dominance*

To examine the effects of language dominance, only data from the bilingual participant groups were included. For cued recall with receptive retrieval (Swahili cues), the fixed factors included language group, the within-subjects factors of response language and recall cycle, and all possible interactions among the factors. The model included random intercepts for participants and items, random slopes for response language and recall cycle across participants and items, and random slopes for language group across items. On this task, the overall accuracy of the two language groups did not differ ( $p > .5$ ). Overall, cued recall was more accurate with English than with Spanish responses,  $b = .326$ ,  $SE = .107$ ,  $z = 3.050$ ,  $p = .002$ . As shown in Figure 2, these factors interacted,  $b = .632$ ,  $SE = .195$ ,  $z = 3.248$ ,  $p = .001$ , such that English-dominant bilinguals had more accurate recall in English than Spanish,  $b = .642$ ,  $SE = .145$ ,  $z = 2.523$ ,  $p < .001$ , but Spanish-dominant participants did not show this effect ( $p > .5$ ). Accuracy improved across the three study-test cycles, as indicated by a significant linear trend,  $b = 2.323$ ,  $SE = .078$ ,  $z = 29.659$ ,  $p < .001$ . Higher cognitive assessment scores were associated with greater accuracy,  $b = .020$ ,  $SE = .010$ ,  $z = 1.979$ ,  $p = .048$  (see Table B.3 for additional details.)

For cued recall with productive retrieval (Swahili responses), the fixed factors were language group, the within-subjects factor of cue language, and their interaction. The model included random intercepts for participants and items, random slopes for cue language across participants and items, and random slopes for group and the interaction of group and cue language across items. Overall accuracy of the two groups did not differ,

( $p > .5$ ). Cued recall was more accurate overall with English than with Spanish cues,  $b = .508$ ,  $SE = .105$ ,  $z = 4.858$ ,  $p < .001$ . These factors interacted,  $b = .550$ ,  $SE = .188$ ,  $z = 2.927$ ,  $p = .003$ , such that English-dominant bilinguals had more accurate recall with English cues than with Spanish cues,  $b = .783$ ,  $SE = .146$ ,  $z = 5.350$ ,  $p < .001$ , but Spanish-dominant participants did not show this effect,  $b = .233$ ,  $SE = .135$ ,  $z = 1.734$ ,  $p = .083$ . Cognitive scores were not significantly associated with accuracy ( $p > .1$ ) (see Table B.4 for additional details.)

For associative recognition, the  $d'$  scores (illustrated in Figure 3) were submitted to a 2 (bilingual group) x 2 (known word language) mixed ANCOVA with cognitive assessment scores as a covariate. Bilingual associative recognition was not significantly more accurate in English than in Spanish,  $F(1, 93) = 1.513$ ,  $MSE = .447$ ,  $p = .222$ ,  $\eta_p^2 = .016$ , and there was no main effect of bilingual group, ( $F < 1$ ). However, there was a significant interaction of bilingual group and response language,  $F(1, 93) = 5.097$ ,  $MSE = .447$ ,  $p = .026$ ,  $\eta_p^2 = .052$ , such that the English-dominant group had an advantage for English over Spanish,  $F(1, 47) = 8.913$ ,  $MSE = .368$ ,  $p = .004$ ,  $\eta_p^2 = .159$ , but the Spanish-dominant group did not,  $F(1, 47) = .160$ ,  $MSE = .532$ ,  $p = .691$ ,  $\eta_p^2 = .003$ . Higher cognitive assessment scores were associated with greater accuracy,  $F(1, 93) = 4.333$ ,  $MSE = 2.089$ ,  $p = .040$ ,  $\eta_p^2 = .045$ .

### Language proficiency

Simple correlations of proficiency scores with cued recall and associative recognition accuracy are given in Table 3. The two bilingual groups were pooled to make the proficiency scores closer to normally distributed. For recall with Swahili cues, the average recall score across blocks was used. A Bonferroni correction was applied to each set of six correlations with a familywise error rate of .05 ( $\alpha = .0083$ ). In both monolingual and bilingual participants, English proficiency scores were significantly correlated with all measures from the English block. For bilinguals, Spanish proficiency scores were significantly associated with all measures

from the Spanish block. No correlations of Spanish proficiency scores with English performance or English proficiency scores with Spanish performance survived the Bonferroni correction. However, the nonsignificant correlations are not so low as to provide strong evidence that only proficiency in the involved known language is associated with learning. Note that proficiency scores in English and Spanish were not correlated in bilingual participants,  $r(94) = -.026$ ,  $p = .803$ , or in monolingual participants,  $r(46) = -.153$ ,  $p = .300$ . Therefore, any positive associations between language proficiency in the uninvolved known language and accuracy cannot be an artifact of such a correlation.

For more rigorous inferential assessments of the associations of proficiency with cued recall, we conducted logistic mixed-effects regression analyses with proficiency scores (W scores on the WMLS-R) as a covariate predictor. For cued recall with receptive retrieval (Swahili cues), the fixed factors were response language, recall cycle (within-subjects), and their interaction, and proficiency scores in the response language was the covariate. Note that this means that English proficiency scores were entered for English response trials, and Spanish proficiency scores were entered for Spanish response trials. The model included random intercepts for participants and items and random slopes for response language and recall cycle across participants and items. Higher proficiency scores in the response language were associated with greater accuracy in receptive retrieval,  $b = .024$ ,  $SE = .004$ ,  $z = 5.354$ ,  $p < .001$  (see Table B.5 for detailed results.)

For cued recall with productive retrieval (Swahili responses), the fixed factor was cue language, and the covariate was proficiency scores in the cue language. The model included random intercepts for participants and items and random slopes for cue language across both participants and items. As expected, higher proficiency scores in the cue language were associated with greater accuracy in productive retrieval,  $b = .024$ ,  $SE = .004$ ,  $z = 5.620$ ,  $p < .001$  (see Table B.6 for detailed results.)

We also conducted proficiency analyses within the monolingual participant group. For receptive retrieval, the fixed factor was recall cycle (within-subjects), and the covariate was English proficiency scores. The model included random intercepts for items and participants and random slopes for recall cycle across participants and items. Participants with higher English proficiency scores had more accurate receptive retrieval,  $b = .024$ ,  $SE = .011$ ,  $z = 2.284$ ,  $p = .022$ . For productive retrieval, the only predictor was the covariate English proficiency scores, and the model included random intercepts for items and participants. Participants with higher English proficiency scores had more accurate productive retrieval,  $b = .069$ ,  $SE = .023$ ,  $z = 3.044$ ,  $p = .002$ . When these analyses were conducted using Spanish proficiency scores, no significant associations were observed ( $ps > .2$ ) (see Tables B.7, B.8, B.9, and B.10 for detailed results.)

### Strategy use

We compared self-reported learning strategies across language groups. Responses to the strategy use questionnaires were coded according to several possible categories, mostly drawn from previous research (Bangert & Heydarian, 2017; Tsuboi & Francis, 2020). The proportion of participants who reported each strategy is given in Table 4 for each language group (note that 2 monolingual participants did not complete the strategy questionnaire.) Reported study strategies were coded into the following categories: (1) word association or mediator, (2) sentence or phrase formation, (3) imagery, (4) rote repetition, (5) structural similarities,

**Table 3** Correlations of Language Proficiency and Associative Memory Measures

Language Group	Cued Recall		
	Swahili Cue (Overall)	Swahili Response	Associative Recognition
English Monolinguals ( $n = 48$ )			
English Blocks			
English proficiency	.377**	.406**	.432**
Spanish proficiency	.132	.088	.128
Spanish-English Bilinguals ( $n = 96$ )			
English Block			
English proficiency	.367***	.369***	.224*
Spanish proficiency	.265*	.204*	.111
Spanish Block			
English proficiency	.177	.150	.110
Spanish proficiency	.483***	.439***	.301**

Note. Swahili Cue (Overall) refers to the mean of the three recall cycles with Swahili cues. Associative recognition correlations are with  $d'$  scores. Proficiency scores are Oral Language composite (calibrated W) scores from the WMLS-R (Woodcock et al., 2005)  
\* $p < .05$ , \*\* $p < .0083$ , \*\*\* $p < .001$



**Table 4** Self-Reported Strategy Use (Proportion)

Reported Strategy	English Monolingual	English Dominant	Spanish Dominant	Percent Agreement	Cohen's $\kappa$
Word Association or Mediator	.61	.58	.56	.84	.68
Sentence or Phrase Formation	.15	.23	.13	.95	.79
Imagery	.11	.35	.29	.96	.89
<i>At Least One Associative Strategy<sup>a</sup></i>	.72	.77	.73	.92	.79
Rote Repetition	.13	.15	.06	.93	.60
Structural Similarities	.26	.40	.38	.88	.73
Focus on Spelling	.17	.23	.23	.90	.66
Focus on Pronunciation	.30	.33	.42	.94	.87
None of the above strategies	.04	.04	.06	--	--

Note. Each proportion listed for individual strategy categories includes participants who reported multiple strategies.

<sup>a</sup>Proportion of participants who reported at least one of the three associative learning strategies listed above.

(6) focus on spelling, and (7) focus on pronunciation. Individual responses were coded into as many categories as were reported. Inter-rater agreement (proportion consistent) and reliability (Cohen's  $\kappa$  – Xonev et al., 1993) were computed for each category (see Table 4). Two raters coded strategies for 22 participants together and then independently coded strategies for the remaining 120 participants (2 participants did not complete the strategy questionnaire). All disagreements were resolved by consensus. We considered the word association or mediator, sentence or phrase formation, and imagery strategies to be associative in nature, and the proportion of participants who reported at least one of these associative strategies was computed. As shown in Table 4, monolingual and bilingual participants used at least one associative strategy at similar rates,  $X^2(1, N = 142) = .172, p = .679, \phi = .035$ . Bilinguals were more likely than monolinguals to report using imagery,  $X^2(1, N = 142) = 7.541, p = .006, \phi = .230$ , but no language-group differences were observed for the other strategies ( $ps > .1$ ).

To examine whether the use of associative strategies was associated with higher accuracy on the memory measures, a 3 (language group)  $\times$  2 (use of associative strategies) ANOVA was performed for each of the six memory measures. To correct for multiple comparisons, a Bonferroni correction with a familywise error rate of .05 was applied ( $\alpha = .0083$ ). Participants who reported using at least one associative strategy had higher accuracy on cued recall with Swahili cues and Spanish responses than those who did not,  $F(1, 94) = 8.701, MSE = .034, p = .004, \eta_p^2 = .085$ . However, none of the other measures yielded a significant associative strategy effect ( $ps > .05$ ). Reported use of imagery in particular also was not associated with higher accuracy on any of the learning measures ( $ps > .05$ ).

## Discussion

The first goal of the study was to rigorously compare monolingual and bilingual learning of vocabulary in an unfamiliar language and explore the plausibility of strategic factors as a mechanism for any group differences. When Swahili cues were used to retrieve English words and in associative recognition, there were no significant differences in accuracy between English monolingual and English-dominant bilingual performance, consistent with our previous research (Tsuboi & Francis, 2020) and in contrast

to other previous research (e.g., Bogulski et al., 2018; Kaushanskaya & Marian, 2009). If anything, the monolingual scores were numerically higher. We had predicted that English-dominant bilinguals would outperform English-speaking monolinguals in the more difficult productive retrieval task in which English words were used to cue the unfamiliar Swahili words. However, the opposite pattern was observed, suggesting that bilinguals did not have an advantage based on their larger total vocabulary (e.g., Gollan et al., 2008) or greater phonological long-term knowledge (e.g., Kaushanskaya, 2012). Across learning measures involving English, English monolingual performance was more accurate than Spanish-dominant bilingual performance. Based on the bilingual results to follow, we interpret these differences as proficiency-related effects.

Note that the three groups were well matched on age, education, and parent education level. If anything, the English monolingual group scored lower on the measure of nonverbal cognitive ability, so there is no evidence that the bilinguals in the present study were cognitively disadvantaged. We intentionally did not match groups on language proficiency scores, because bilinguals on average have lower vocabulary in each of their languages due to splitting their usage between languages (e.g., Gollan et al., 2008). Therefore, matching monolingual and bilingual vocabulary scores would result in non-representative bilingual samples with particularly high language-learning ability (as pointed out by Prior & MacWhinney, 2010; Tsuboi & Francis, 2020) or non-representative monolingual samples with particularly low language-learning ability.

We examined whether the reported use of associative strategies differed across groups and whether it was predictive of learning. Bilingual and monolingual participants reported the use of associative strategies for learning the word pairs at similar rates (about 74%), although bilinguals reported the use of imagery more frequently than monolinguals. The use of associative strategies was not consistently associated with more efficient learning. This result converges with previous findings that self-reported use of particular strategies was not associated with cued recall of words in an unfamiliar language with productive retrieval (Bangert & Heydarian, 2017; Tsuboi & Francis, 2020).

The second goal of the study was to clarify the relationships of language dominance and proficiency in known languages to the efficiency of learning vocabulary in an unfamiliar language.

The three learning measures, recall with Swahili cues, recall with Swahili responses, and associative recognition showed consistent patterns of performance with respect to language dominance group and task language. In each case, there was a significant interaction of language dominance group with the nominal language of the task, such that performance was more accurate when the task was performed in the dominant language. Overall, accuracy was higher in English than in Spanish. Given that the degree of dominance of the two bilingual groups was similar (as shown by WMLS-R scores in Table 1), this effect indicates simply that the task was more difficult in Spanish. Thus, it is a property of the materials, not a characteristic of the participants.

In bilinguals, proficiency in English or Spanish was correlated with learning measures that involved the same language (with the exception of associative recognition in English), but proficiency was not reliably correlated with learning measures in the other language (a pattern consistent with Tsuboi & Francis, 2020). Contrary to our previous study (Tsuboi & Francis, 2020), monolingual English proficiency was also significantly correlated with learning measures. Note, however, that Spanish proficiency scores in monolingual participants were not correlated with the learning measures. Therefore, the absence of a bilingual advantage cannot be explained by the monolingual participants' knowledge of Spanish vocabulary.

The language-dominance effects and proficiency correlations extended to the more difficult cued recall task that required productive retrieval, where known words were given as cues to recall the unfamiliar Swahili words. These correlations were of comparable strength to those observed for cued recall with receptive retrieval. Therefore, the proficiency effects observed cannot be attributed to the relative accessibility of the known words. Instead, the locus of the effects of proficiency must be in the strength of the new associations formed at encoding. In the following section we consider possible explanations for

why higher proficiency would be associated with the formation of stronger associations.

### Possible mechanisms for proficiency effects

We considered several possible mechanisms for the observed associations between language proficiency and the learning measures. First, we evaluated explanations that involve a causal relationship between language proficiency and the processes of learning. One possibility was that the higher vocabulary of more proficient speakers would give them more possible mediator words. If this were based on total vocabulary (summed across languages), we would have expected more efficient learning in bilinguals. In contrast, if the important factor were vocabulary in the specific language with which the novel Swahili words were presented, we would have expected consistently more efficient learning in monolinguals. A related possibility was that because higher proficiency is associated with stronger links between words and their concepts, mediators could be identified more rapidly, but this explanation would be expected to favor monolingual participants.

We also considered measured third variables that might underlie both proficiency and learning of vocabulary in a new language, including cognitive ability, socio-economic status, experience with the known language, and language learning ability, and relevant correlations are given in Table 5. A Bonferroni correction with a familywise error rate of .05 was applied for each predictor across the six learning measures ( $\alpha = .0083$ ). With respect to cognitive ability, non-verbal cognitive ability scores (WJ-III visuo-spatial ability) were not correlated with any of the learning measures in monolinguals or bilinguals. Socio-economic status scores (parent education scores) were not correlated with the learning measures in monolinguals or bilinguals. For bilinguals, experience with English was estimated by subtracting the age of acquisition from the age of the participant,

**Table 5** Correlations of Cognitive, Demographic, and Experiential Variables with Memory Measures

Language Group	English Cued Recall		English	Spanish Cued Recall		Spanish
	Swahili Cue (Overall)	Swahili Response	Associative Recognition	Swahili Cue (Overall)	Swahili Response	Associative Recognition
English Monolinguals ( $n = 48$ )						
Visuo-Spatial Reasoning <sup>a</sup>	-.068	-.029	-.033	-	-	-
Parent Education <sup>b</sup>	.128	.151	.161	-	-	-
Spanish-English Bilinguals ( $n = 96$ )						
Visuo-Spatial Reasoning <sup>a</sup>	.212*	.155	.264*	.174	.114	.135
Parent Education <sup>b</sup>	.012	-.049	.051	.132	.019	.214*
Experience with English <sup>c</sup>	-.045	-.036	-.144	-.002	-.088	-.062
English Proficiency Controlling for Experience with English <sup>c</sup>	.475***	.456***	.339***	.218*	.232*	.169
Spanish Proficiency Controlling for Experience with English <sup>c</sup>	.276**	.208*	.084	.505***	.440***	.301**

Note. Swahili Cue (Overall) refers to the mean of the three recall cycles with Swahili cues.

<sup>a</sup>W scores from the Woodcock-Johnson III Test of Cognitive Abilities (Woodcock McGrew, & Mather, 2001)

<sup>b</sup>Parent education scored on 1-6 scale (1 = less than 8<sup>th</sup> grade; 2 = some high school; 3 = graduated high school; 4 = some college; 5 = graduated college; 6 = advanced degree); data were missing for 3 monolingual participants.

<sup>c</sup>Experience with English obtained by subtracting age of English acquisition from age of participant; English experience was missing for 1 bilingual participant.

\* $p < .05$ , \*\* $p < .0083$ , \*\*\* $p < .001$

and these scores also were not correlated with the learning measures.

To examine the possible influence of language learning skill, we computed the partial correlation between English proficiency scores and English learning measures in bilinguals, controlling for years of English experience. These positive correlations were highly significant for all three learning measures. Similarly, the partial correlations between Spanish proficiency scores and the Spanish learning measures controlling for years of English experience were significant. These partial correlations suggest that higher language-learning skill underlies both higher proficiency attainment in known languages and more efficient learning on the new task. While this relationship is not particularly surprising, the mechanisms by which higher language-learning skill might lead to more efficient learning of translation paired associates is not obvious. That is, this relationship alone does not tell us what individuals with higher language learning skill do differently when learning the translation pairs. One possibility was a greater reliance on associative strategy use. Indeed, when controlling for years of English experience, bilinguals with higher English proficiency scores were more likely to report using associative strategies,  $r(92) = .248$ ,  $p = .016$ . However, we did not obtain reliable evidence that the use of associative strategies enhanced learning.

We also considered the possible role of phonological working memory, although we had no direct measures of this construct. Previous research has shown that in both monolingual and bilingual children, better phonological working memory is associated with more extensive vocabulary in their first-learned languages (e.g., Gathercole et al., 1999; Lanfranchi & Swanson, 2005; Swanson et al., 2011) and later-learned languages (Lanfranchi & Swanson, 2005; Swanson et al., 2011). In young adulthood these patterns persist for bilinguals (Kaushanskaya et al., 2011), but are less consistent for monolinguals (e.g., Kaushanskaya et al., 2011; Kemper & Sumner, 2001), perhaps because other factors have a greater influence (Gathercole et al., 1999). Learning vocabulary in an unfamiliar language is more efficient for monolingual (Kaushanskaya, 2012) and bilingual/multilingual adults (Papagno & Vallar, 1995) with higher phonological working memory. Thus, given that phonological working memory is associated with more extensive vocabulary in known languages and more efficient learning of vocabulary in unfamiliar languages, more extensive vocabulary should be associated with more efficient learning of vocabulary in unfamiliar languages. Indeed, this association was evident in both the bilingual and monolingual samples. Thus, higher phonological working memory is a plausible causal factor underlying this association.

### External validity considerations

It is important to note that the bilingual participants in the present study generally learned English out of necessity while living and going to school in the U.S. Thus, their bilingualism did not arise out of a particular interest in or ability to learn a second language, nor were they participants in an enriched education program targeting students with high academic performance or high socio-economic status. Thus, the present results are not expected to generalize to conditions in which learning a second language is based on self-selection or targeted educational opportunities, where bilinguals might well be expected to outperform monolinguals on the paired-associate learning task.

Because the monolingual participants lived in environment where they could be exposed to Spanish on a regular basis, we

considered whether monolingual participants with more knowledge of Spanish might have learned more efficiently, perhaps masking a bilingual advantage. However, in the monolingual group, Spanish proficiency scores were not correlated with performance on the learning tasks (consistent with Tsuboi & Francis, 2020). Also, in a previous study with monolingual participants, ambient language diversity was not associated with enhanced learning of vocabulary in a new language (Bice & Kroll, 2019).

With respect to generalizing the results of the present study, it should be noted that vocabulary acquisition is only one of many important aspects of learning a new language. Thus, the present results do not have implications for other critical components of language learning. Also, studying translation pairs is not the only way to learn vocabulary in a new language. Note, however, that the advantage for the dominant language and the proficiency effects in the present study are consistent with the results of a previous study in which unfamiliar words in a known language were learned by inferring meanings from sentence contexts (Lauro et al., 2020).

Because none of the stimuli were cognates or non-cognate homographs, which make learning easier (de Groot & Keijzer, 2000), bilinguals may not have been able to take full advantage of their bilingual experience and knowledge. Knowing two or more languages increases the probability that words in a new language will have cognates in a known language, and a bilingual person would likely take advantage of such similarities. Therefore, in more natural learning contexts, we might indeed expect bilingual-monolingual learning differences.

The small number of available normed stimuli meant that the experiment could not be designed to directly compare independent measures of receptive and productive retrieval within participants along with the language manipulation, and it is possible that the three cycles of receptive retrieval practice influenced the pattern of later productive retrieval and/or associative recognition across groups.

Finally, learning was tested using a relatively short retention interval (within a single experimental session), and it is uncertain whether the results would generalize to the longer retention intervals associated with more natural learning contexts. Indeed, factors that affect verbal learning performance at shorter and longer intervals are not always the same (e.g., Schmidt & Bjork, 1992).

### Conclusions

Bilinguals learning vocabulary in an unfamiliar language through their more proficient language did not outperform matched monolinguals, even in a difficult task in which they had to respond in the unfamiliar language. Thus, neither a larger total vocabulary nor greater phonological long-term knowledge conferred a benefit to the bilingual participants. The reported use of associative strategies was similar for monolingual and bilingual participants and did not reliably predict learning. Bilinguals learned vocabulary more efficiently through their more proficient language, as indicated by more accurate performance on all learning measures. Higher proficiency in the known language involved in the learning task was associated with higher accuracy even when recalling the unfamiliar Swahili words. Thus, in this learning context, higher proficiency is associated with more efficient or stronger learning of associations rather than more efficient retrieval of the known words. The association between proficiency

in known languages and performance on the learning measures may arise from variation in language learning skill, for which phonological working memory is a plausible mechanism of influence.

**Data availability.** The data that support these findings are available at <https://data.mendeley.com/datasets/wbwsk44pwb/1>.

**Acknowledgements.** This research was supported in part by NSF Grant 1632283 to Francis and by research internships awarded to Nájera, including RISE (funded by NIH Grant 2R25GM069621) and SURPASS (funded jointly by the UTEP Vice President of Student Affairs and Campus Office for Undergraduate Research Initiatives). We gratefully acknowledge the assistance of Naoko Tsuboi in implementation and training of research assistants and the assistance of Rebekah Villaverde, Luis Urrea, Isabela Blanco Almeida, Ileen Gurrola-Meza, and Kimberly Rivera in data collection and data entry.

**Competing Interests.** There are no competing interests to declare.

## References

- Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, 6, 451-474.
- Bangert, A. S., & Heydarian, N. M. (2017). Recall and response time norms for English-Swahili word pairs and facts about Kenya. *Behavior Research Methods*, 49, 124-171.
- Bice, K., & Kroll, J. F. (2019). English only? Monolinguals in linguistically diverse contexts have an edge in language learning. *Brain & Language*, 196, Article 104644.
- Bogulski, C. A., Bice, K., & Kroll, J. F. (2018). Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. *Bilingualism: Language and Cognition*, 22(5), 1052-1067.
- Clark, S. E. (1992). Word frequency effects in associative and item recognition. *Memory & Cognition*, 20, 231-243.
- 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(2), 257-271.
- de Groot, A. M., & Keijzer, R. (2000). What Is Hard to Learn Is Easy to Forget: The Roles of Word Concreteness, Cognate Status, and Word Frequency in Foreign-Language Vocabulary Learning and Forgetting. *Language Learning*, 50(1), 1-56.
- Galobardes, B., Shaw, M., Lawlar, D. A., Lynch, J. W., & Davey Smith, G. (2006). Indicators of socioeconomic position (part 1). *Journal of Epidemiological Community Health*, 60, 7-12.
- Gathercole, S. E., Service, E., Hitch, G. J., Adams, A. M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology*, 13, 65-77.
- 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(3), 787-814.
- Griffin, G., & Harley, T. (1996). List learning of second language vocabulary. *Applied Psycholinguistics*, 17(4), 443-460.
- The jamovi project. (2022). Jamovi. (Version 2.3) [Computer Software]. Retrieved from <https://www.jamovi.org>.
- Kaushanskaya, M. (2012). Cognitive mechanisms of word learning in bilingual and monolingual adults: The role of phonological memory. *Bilingualism: Language and Cognition*, 15, 470-489.
- Kaushanskaya, M., Blumenfeld, H. K., & Marian, V. (2011). The relationship between vocabulary and short-term memory measures in monolingual and bilingual speakers. *International Journal of Bilingualism*, 15, 408-425.
- Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, 16(4), 705-710.
- Kaushanskaya, M., & Reetzgel, K. (2012). Concreteness effects in bilingual and monolingual word learning. *Psychonomic Bulletin & Review*, 19(5), 935-941.
- Kemper, S., & Sumner, A. (2001). The structure of verbal abilities in young and older adults. *Psychology and Aging*, 16, 312-322.
- Lanfranchi, S., & Swanson, H. L. (2005). Short-term memory and working memory in children as a function of language-specific knowledge in English and Spanish. *Learning and Individual Differences*, 15, 299-319.
- Lauro, J., Schwartz, A. I., & Francis, W. S. (2020). Bilingual novel word learning in sentence context. *Journal of Memory and Language*, 113, 104123.
- Madan, C. R., Glaholt, M. G., & Caplan, J. B. (2010). The influence of item properties on association-memory. *Journal of Memory and Language*, 63, 46-63.
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, 87, 252-271.
- Nelson, T. O., & Dunlosky, J. (1994). Norms of paired-associate recall during multitrial learning of Swahili-English translation equivalents. *Memory*, 2, 325-335.
- Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology*, 48A, 98-107.
- Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, 13, 253-262.
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, 3, 207-218.
- Schneider, V. I., Healy, A. F., & Bourne, L. E. (2002). What is learned under difficult conditions is hard to forget: Contextual interference effects in foreign vocabulary acquisition, retention, and transfer. *Journal of Memory and Language*, 46(2), 419-440.
- Schrank, F. A., McGrew, K. S., & Dailey, D. E. H. (2010). *Technical Supplement. Woodcock-Muñoz Language Survey-Revised Normative Update*. Rolling Meadows, IL: Riverside Publishing.
- Swanson, H. L., Orosco, M. J., & Lussier, C. M. (2011). The influence of working memory and phonological processing on English language learner children's bilingual reading and language acquisition. *Journal of Educational Psychology*, 103, 838-856.
- Tsuboi, N., & Francis, W. S. (2020). Rethinking bilingual enhancement and dominance effects in associative learning of foreign language vocabulary: The role of proficiency in the mediating language. *Journal of Memory and Language*, 115, 104155.
- Van Hell, J. G., & Mahn, A. C. (1997). Keyword mnemonics versus rote rehearsal: learning concrete and abstract foreign words by experienced and inexperienced learners. *Language Learning*, 47, 507-546.
- 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., & Alvarado, C. G. (2005). *Woodcock-Muñoz Language Survey-Revised*. Rolling Meadows, IL: Riverside Publishing.

## Appendix A: Word Stimuli

SET A			SET B		
English	Spanish	Swahili	English	Spanish	Swahili
agent	agente	wakili	beer	cerveza	pombe
almond	almendra	lozi	bribe	soborno	rushwa
blame	culpa	lawama	bucket	cubeta	ndoo
boat	barco	mashua	bull	toro	fahali
bone	hueso	mfupa	cinnamon	canela	dalasini
broom	escoba	fagio	cloud	nube	wingu
carpet	alfombra	zulia	corn	elote	mafaka
cheese	queso	jibini	cotton	algodón	pamba
divorce	divorcio	talaka	crab	cangrejo	kaa
dog	perro	mbwa	curtain	cortina	pazia
donkey	burro	punda	custom	costumbre	desturi
dust	polvo	vumbi	egg	huevo	yai
economy	economía	iktisadi	enemy	enemigo	adui
flavor	sabor	ladha	flood	indundar	gharika
forehead	frente	paji	food	comida	chakula
gate	portón	lango	friend	amigo	rafiki
glue	pegamento	ambo	frog	rana	chura
horse	caballo	farasi	garden	jardín	bustani
knee	rodilla	goti	grapes	uvas	zabibu
mattress	colchón	godoro	lake	lago	ziwa
monkey	chango	tumbili	leaf	hoja	jani
oath	juramento	yamini	lung	pulmón	pafu
ornament	ornamento	rembo	merchant	comerciante	tajiri
parrot	perico	kasuku	mystery	misterio	fumbo
poison	veneno	sumu	orphan	huerfano	yatima
prophet	profeta	nabii	pearl	perla	lulu
rope	cuerda	kamba	plate	plato	sahani
sailor	marinero	baharia	poem	poema	utenzi
scarf	bufanda	leso	queen	reina	malkia
science	ciencia	elimu	silk	seda	hariri
snow	nieve	theluji	sleep	sueño	usingizi
society	sociedad	chama	story	cuento	hadithi
treasure	tesoro	dafina	tomato	tomate	nyanya
wheel	rueda	duara	wound	herida	jeraha

## Appendix B: Tables of Logistic Mixed-Effects Regression Results

**Table B.1** Fixed and Random Effects for Receptive Recall in English

Fixed Effects	Estimate	SE	z	p	Odds Ratio (95% CI)
(Intercept)	-0.611	0.161	-3.804	< .001	.543 (.396, .744)
Group (ED - EM)	-0.465	0.273	-1.705	.088	.628 (.368, 1.072)
Group (SD - EM)	-0.773	0.277	-2.796	.005	.461 (.268, .793)
Cycle (linear)	2.408	0.078	30.834	< .001	11.114 (9.537, 12.953)
Cycle (quadratic)	-0.186	0.049	-3.798	< .001	.830 (.754, .914)
Cognitive	0.021	0.009	2.343	.019	1.022 (1.003, 1.040)
Group (ED - EM) x Cycle (linear)	-0.077	0.153	-0.503	.615	.926 (.686, 1.249)
Group (SD - EM) x Cycle (linear)	-0.364	0.152	-2.391	.017	.695 (.516, .936)
Group (ED - EM) x Cycle (quadratic)	-0.034	0.082	-0.411	.681	.967 (.823, 1.136)
Group (SD - EM) x Cycle (quadratic)	-0.079	0.082	-0.957	.338	.924 (.787, 1.086)
Random Effects	Variance				
Participant (intercept)	1.674				
Cycle (linear) (slope)	0.303				
Cycle (quadratic) (slope)	0.001				
Target Word (intercept)	0.913				
Group (ED - EM) (slope)	0.110				
Group (SD - EM) (slope)	0.221				
Cycle (linear) (slope)	0.111				
Cycle (quadratic) (slope)	0.063				
Residual	1.000				

Note. CI = confidence interval, EM = English monolingual, ED = English-dominant bilingual, SD = Spanish-dominant bilingual.

**Table B.2** Fixed and Random Effects for Productive Recall in English

Fixed Effects	Estimate	SE	z	p	Odds Ratio (95% CI)
(Intercept)	0.085	0.182	0.465	.642	1.088 (.762, 1.555)
Group (ED - EM)	-0.658	0.310	-2.125	.034	.518 (.282, .950)
Group (SD - EM)	-0.801	0.314	-2.550	.011	.449 (.242, .831)
Cognitive	0.016	0.015	1.122	.262	1.016 (.988, 1.046)
Random Effects	Variance				
Participant (intercept)	2.096				
Target Word (intercept)	1.179				
Group (ED - EM) (slope)	0.051				
Group (SD - EM) (slope)	0.205				
Residual	1.000				

Note. CI = confidence interval, EM = English monolingual, ED = English-dominant bilingual, SD = Spanish-dominant bilingual.

**Table B.3** Fixed and Random Effects for Receptive Recall in Bilinguals

Fixed Effects	Estimate	SE	z	p	Odds Ratio (95% CI)
(Intercept)	-0.967	0.165	-5.867	< .001	.380 (.275, .525)
Group (ED - SD)	0.000	0.245	0.001	.999	1.000 (.619, 1.615)
BlockLang (E - S)	0.326	0.107	3.050	.002	1.385 (1.123, 1.708)
Cycle (linear)	2.323	0.078	29.659	< .001	10.210 (8.757, 11.905)
Cycle (quadratic)	-0.267	0.047	-5.693	< .001	.766 (.699, .839)
Cognitive	0.020	0.010	1.979	.048	1.020 (1.000, 1.040)
Group x BlockLang	0.632	0.195	3.248	.001	1.881 (1.285, 2.754)
Group x Cycle (linear)	0.163	0.130	1.249	.212	1.177 (.911, 1.519)
Group x Cycle (quadratic)	0.010	0.071	0.136	.892	1.010 (.879, 1.160)
BlockLang x Cycle (linear)	0.064	0.087	0.734	.463	1.066 (.899, 1.263)
BlockLang x Cycle (quadratic)	0.052	0.068	0.763	.445	1.053 (0.922, 1.204)
Group x BlockLang x Cycle (linear)	0.166	0.162	1.025	.305	1.180 (.859, 1.621)
Group x BlockLang x Cycle (quadratic)	0.094	0.135	0.695	.487	1.098 (.843, 1.430)
Random Effects	Variance				
Participant (intercept)	1.348				
BlockLang (slope)	0.730				
Cycle (linear) (slope)	0.233				
Cycle (quadratic) (slope)	0.010				
Target Word (intercept)	0.848				
Group (slope)	0.117				
BlockLang (slope)	0.120				
Cycle (linear) (slope)	0.089				
Cycle (quadratic) (slope)	0.046				
Residual	1.000				

Note: CI = confidence interval, ED = English-dominant bilingual, SD = Spanish-dominant bilingual, E = English, S = Spanish

**Table B.4** Fixed and Random Effects for Productive Recall in Bilinguals

Fixed Effects	Estimate	SE	z	p	Odds Ratio (95% CI)
(Intercept)	-0.421	0.184	-2.286	.022	.657 (.458, .942)
Group (ED - SD)	-0.107	0.267	-0.401	.688	.899 (.533, 1.515)
BlockLang (E - S)	0.508	0.105	4.858	< .001	1.662 (1.354, 2.040)
Cognitive	0.023	0.015	1.546	.122	1.023 (.994, 1.054)
Group x BlockLang	0.550	0.188	2.927	.003	1.733 (1.199, 2.504)
Random Effects	Variance				
Participant (intercept)	1.604				
BlockLang (slope)	0.429				
Target Word (intercept)	1.079				
Group (slope)	0.001				
BlockLang (slope)	0.158				
Group x BlockLang (slope)	0.077				
Residual	1.000				

Note: CI = confidence interval, ED = English-dominant bilingual, SD = Spanish-dominant bilingual, E = English, S = Spanish

**Table B.5** Fixed and Random Effects in Receptive Recall: Bilingual Proficiency

Fixed Effects	Estimate	SE	z	p	Odds Ratio (95% CI)
(Intercept)	-0.972	0.156	-6.223	< .001	.378 (.279, .514)
BlockLang (E - S)	0.281	0.104	2.701	.007	1.324 (1.080, 1.624)
Cycle (linear)	2.292	0.065	35.161	< .001	9.896 (8.709, 11.245)
Cycle (quadratic)	-0.271	0.036	-7.436	< .001	.763 (.710, .819)
Task Language Proficiency	0.024	0.004	5.354	< .001	1.024 (1.015, 1.033)
BlockLang x Cycle (linear)	0.062	0.086	0.723	.470	1.064 (0.899, 1.259)
BlockLang x Cycle (quadratic)	0.043	0.068	0.629	.530	1.043 (.914, 1.191)
Random Effects	Variance				
Participant (intercept)	1.129				
BlockLang (slope)	0.677				
Cycle (linear) (slope)	0.209				
Cycle (quadratic) (slope)	0.007				
Target Word (intercept)	0.818				
BlockLang (slope)	0.117				
Residual	1.000				

Note: CI = confidence interval, E = English, S = Spanish

**Table B.6** Fixed and Random Effects for Productive Recall: Bilingual Proficiency

Fixed Effects	Estimate	SE	z	p	Odds Ratio (95% CI)
(Intercept)	-0.419	0.178	-2.356	.018	.658 (.464, .932)
BlockLang (E - S)	0.468	0.100	4.683	< .001	1.596 (1.312, 1.941)
Task Language Proficiency	0.024	0.004	5.620	< .001	1.024 (1.016, 1.033)
Random Effects	Variance				
Participant (intercept)	1.405				
BlockLang (slope)	0.348				
Target Word (intercept)	1.069				
BlockLang (slope)	0.150				
Residual	1.000				

Note: CI = confidence interval, E = English, S = Spanish



**Table B.7** Fixed and Random Effects in Receptive Recall: Monolingual English Proficiency

Fixed Effects	Estimate	SE	z	p	Odds Ratio (95% CI)
(Intercept)	-0.227	0.235	-0.967	.334	.797 (.503, 1.262)
Cycle (linear)	2.519	0.118	21.321	< .001	12.418 (9.851, 15.653)
Cycle (quadratic)	-0.149	0.063	-2.351	.019	.862 (.761, .976)
English Proficiency	0.024	0.011	2.284	.022	1.024 (1.003, 1.046)
Random Effects	Variance				
Participant (intercept)	0.975				
Cycle (linear) (slope)	0.104				
Cycle (quadratic) (slope)	0.080				
Target Word (intercept)	1.886				
Cycle (linear) (slope)	0.379				
Cycle (quadratic) (slope)	0.001				
Residual	1.000				

Note: CI = confidence interval

**Table B.8** Fixed and Random Effects in Productive Recall: Monolingual English Proficiency

Fixed Effects	Estimate	SE	z	p	Odds Ratio (95% CI)
(Intercept)	0.572	0.286	2.003	.045	1.772 (1.012, 3.103)
English Proficiency	0.069	0.023	3.044	.002	1.071 (1.025, 1.120)
Random Effects	Variance				
Participant (intercept)	1.459				
Target Word (intercept)	2.741				
Residual	1.000				

Note: CI = confidence interval

**Table B.9** Fixed and Random Effects in Receptive Recall: Monolingual Spanish Proficiency

Fixed Effects	Estimate	SE	z	p	Odds Ratio (95% CI)
(Intercept)	-0.218	0.248	-0.879	.380	.804 (.495, 1.308)
Cycle (linear)	2.528	0.121	20.944	< .001	12.525 (9.887, 15.867)
Cycle (quadratic)	-0.141	0.063	-2.249	.025	.868 (.768, .982)
Spanish Proficiency	-0.004	0.005	-0.900	.368	.996 (.986, 1.005)
Random Effects	Variance				
Participant (intercept)	0.976				
Cycle (linear) (slope)	0.105				
Cycle (quadratic) (slope)	0.079				
Target Word (intercept)	2.198				
Cycle (linear) (slope)	0.409				
Cycle (quadratic) (slope)	0.001				
Residual	1.000				

Note: CI = confidence interval

**Table B.10** Fixed and Random Effects in Productive Recall: Monolingual Spanish Proficiency

Fixed Effects	Estimate	SE	z	p	Odds Ratio (95% CI)
(Intercept)	0.573	0.304	1.886	.059	1.774 (.978, 3.217)
English Proficiency	0.009	0.011	0.857	.391	1.010 (.988, 1.032)
Random Effects	Variance				
Participant (intercept)	1.458				
Target Word (intercept)	3.253				
Residual	1.000				

Note: CI = confidence interval