Applied Psycholinguistics (2023), **44**, 179–203
doi:10.1017/S0142716423000140

CAMBRIDGE UNIVERSITY PRESS

ORIGINAL ARTICLE

The effect of memory instructions on within- and between-language false memory

Maria Soledad Beato[1,*], Pedro B. Albuquerque[2], Sara Cadavid[3] and Mar Suarez[1]

[1]University of Salamanca, Salamanca, Spain, [2]University of Minho, Braga, Portugal and [3]School of Medicine and Health Sciences, Universidad del Rosario, Rosario, Colombia
*Corresponding author. Email: msol@usal.es

(Received 20 May 2022; revised 8 February 2023; accepted 8 March 2023; first published online 29 March 2023)

Abstract

We examined the effect of memory instructions on false memory using the Deese/Roediger–McDermott paradigm in second-language learners. Participants studied lists of words in L1 and L2 (e.g., note, sound, piano . . .) associatively related to a non-presented critical lure (e.g., MUSIC). In a later recognition test, critical lures appeared in the same or the other language of their lists (i.e., within- and between-language conditions). In Experiment 1, participants should only endorse an item when study and test languages matched (i.e., restrictive instructions); that is, they should retrieve language information. In Experiment 2, participants should endorse studied items regardless of the language (i.e., inclusive instructions). With restrictive instructions, false recognition was higher in within- than between-language conditions, whereas with inclusive instructions, this result was replicated only when words were studied in L1, but not L2. Results suggested that second-language learners show false memory in their L2 and that the effect of language shift on false recognition depended on the study language.

Keywords: false memory; DRM paradigm; language shift; second-language learners; memory instructions

The last two decades have witnessed a growing interest in bilingualism. As the world becomes more globalized, the number of bilinguals increases to the point that today there are more bilingual than monolingual speakers (Bialystok, 2017). In this globalized world, not only bilingualism has increased but also the number of people with a certain level of second-language proficiency. It is increasingly common for people to learn and to use a second language at some point in their lives. This context has favored research on the effects of learning a second language on general human functioning (for reviews, see Klimova, 2018; Lehtonen et al., 2018; Quinteros Baumgart & Billick, 2018).

Paralleling the increasing attention to multilingualism phenomena, in the last decades, the study of false memories has attracted the scientific community's efforts, as it constitutes a means of providing insight into the constructive nature of memory (Schacter & Slotnick, 2004; for reviews, see Brainerd & Reyna, 2005; Gallo, 2006, 2010).

Several experimental procedures have been used to induce memory distortions, but one, in particular, has been widely employed: the Deese/Roediger–McDermott (DRM) paradigm (Deese, 1959; Roediger & McDermott, 1995). In this paradigm, participants study lists of words associated with a non-presented critical lure. In a later memory test, not only studied words are retrieved (true memories) but also critical lures are often falsely recalled or recognized, leading to the false memory effect (e.g., Arndt, 2012; Cadavid & Beato, 2017; Wang et al., 2018; Yonelinas et al., 2010).

The DRM paradigm was created in an English-speaking environment, but the robustness of the false memory effect has fascinated memory researchers worldwide (e.g., Arndt, 2015; Beato & Arndt, 2017; Beato et al., 2023; Cadavid et al., 2021; Huff et al., 2014; Wang et al., 2019). Consequently, DRM materials have been used in different languages by applying two different procedures. On the one hand, researchers have created new DRM lists based on free association norms in their own language to study false memories (e.g., Chinese: Chen et al., 2008; Geng et al., 2007; Lee et al., 2007; Dutch: Van Damme & D'Ydewalle, 2009a, 2009b; French: Dubuisson et al., 2012; Plancher et al., 2008; Hebrew: Anaki et al., 2005; Ben-Artzi et al., 2009; Italian: Iacullo & Marucci, 2016; Japanese: Abe et al., 2008; Kawasaki & Yama, 2006; Polish: Ulatowska & Olszewska, 2013; Brazilian Portuguese: Stein et al., 2006; European Portuguese: Albuquerque, 2005; Albuquerque & Pimentel, 2005; Rocha & Albuquerque, 2003; European Spanish: Beato & Arndt, 2014; Boldini et al., 2013; Mexican Spanish: Anastasi, De Leon et al., 2005; Swedish: Johansson & Stenberg, 2002). On the other hand, to study the false memory effect, researchers have also translated original English lists (Roediger & McDermott, 1995; Stadler et al., 1999) to a second language, such as Chinese (Mao et al., 2010), Dutch (Zeelenberg & Pecher, 2002), French (Cabeza & Lennartson, 2005; Howe et al., 2008), German (Diekelmann et al., 2008, 2010; Rummer et al., 2009), Italian (Ciaramelli et al., 2006), Brazilian Portuguese (Stein & Pergher, 2001), or European Spanish (García-Bajos & Migueles, 1997) among others. In this regard, although a direct translation does not seem to be the best practice to adapt DRM materials (for reviews of methodological issues, see Graves & Altarriba, 2014; Marmolejo et al., 2009), robust false memory effects have been reported both translating original English lists to a second language and creating new lists based on free association norms of the language of interest.

One of the main theories that account for the DRM false memory illusion is the Activation-Monitoring Framework or AMF (Roediger et al., 2001). The AMF assumes that two different memory processes work in opposition to create a false memory: activation and monitoring. Specifically, false memory occurs when, first, the critical lure is automatically activated due to preexisting associations between the studied words and the critical lure, and, subsequently, monitoring processes fail. Another prominent theory of DRM false memories is the Fuzzy-Trace Theory or FTT (Brainerd & Reyna, 2002),[1] which also posits the existence of two opposing processes in false memory formation. In this framework, false memory appears when the gist information of the list is extracted, and this gist trace matches the critical lure, and the retrieval of verbatim representations is not enough to reject the critical lure (i.e., no recollection rejection).

Several models have also been proposed to explain how second languages are represented in the brain. Currently, the most accepted models are the Revised

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

Hierarchical Model (Kroll & Stewart, 1994; Kroll et al., 2010) and the connectionist models of bilingual representation (Hernandez et al., 2005). These models differ on whether second-language lexical entries are developed and stored separately (Kroll & Stewart, 1994) or not (Hernandez et al., 2005). However, they agree on a critical characteristic: first-language lexical entries are quicker to fully activate conceptual representations or concepts than second-language lexical entries. Previous research pointed at this shared characteristic to establish straightforward predictions on false memory effects across first or dominant languages (L1, hereafter) and second or non-dominant languages (L2, hereafter) (e.g., Arndt & Beato, 2017; Beato & Arndt, 2021). Specifically, false memory is expected to be higher in the L1 than L2 because there would be a more automatic access to the conceptual representations from the dominant language. This effect is known as the language dominance effect. Furthermore, according to the Revised Hierarchical Model, second-language learners or unbalanced bilinguals access concepts from L2 words through their L1 translation. That is, participants with low L2 proficiency would not directly access the concept from L2 words but need to translate that word into their L1 and subsequently access the conceptual representation.

After reviewing the literature, we found that, despite the aforementioned growing interest in both false memory and bilingualism, false memory in a non-dominant language is still a poorly studied and understood phenomenon. Besides this, the few studies that have investigated false memory in both the L1 and L2 with the DRM paradigm have used different experimental conditions, reaching inconsistent conclusions. The goal of the current research is to address this gap. We studied memory processes in the L1 and L2 focusing on the role of automatic associations, so we need to bridge the false memory and language literature traditions. Concerning false memory literature, in this particular research, both the AMF and the FTT would make the same predictions because the associative strength between the studied words and the critical lure correlates with processes that allow gist extraction (Cann et al., 2011; Huff et al., 2021). Regarding the language literature, the Revised Hierarchical Model (RHM) and the connectionist models rely on associative activations, just like the AMF.

It is important to understand how false memories are generated not only in dominant but also in non-dominant languages because it allows us, first, to establish the nature of the false memory phenomenon in the L1 and L2, and second, to explore the role that automatic access to conceptual representations plays in false memory. The next section presents an exhaustive review of research focused on false memory in the L1 and L2 in order to understand why studies sometimes reached different conclusions and how to move forward in this field.

Previous findings in within- and between-language false memory

An exhaustive literature review on false memory revealed that different experimental conditions have been used, with some studies including conditions in which study and test language matched (i.e., within-language conditions, L1L1 and L2L2) (Anastasi, Rhodes, et al., 2005; Arndt & Beato, 2017), and others including conditions in which study and test languages did not match (i.e., between-language conditions, L1L2 and L2L1) (Cabeza & Lennartson, 2005; Howe et al., 2008;

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

Kawasaki-Miyaji et al., 2003; Marmolejo et al., 2009; Sahlin et al., 2005; see Graves & Altarriba for a review).

Firstly, regarding within-language studies, the results showed that highly proficient participants in the L1, with less proficiency in the L2, systematically produced higher false recognition when words were studied in the L1 (i.e., L1L1) than in the L2 (i.e., L2L2) (Anastasi, Rhodes, et al., 2005, Experiments 3 and 4; Arndt & Beato, 2017; Beato & Arndt, 2021; Howe et al., 2008; Kawasaki-Miyaji et al., 2003; Marmolejo et al., 2009; Sahlin et al., 2005). Only when participants were highly exposed to their L2 (they lived in an L2 environment), researchers found higher levels of memory distortion in the L2 than in the L1 (Anastasi, Rhodes, et al., 2005, Experiment 2). For its part, when L1 and L2 proficiency was similar, false recognition did not differ between languages (Cabeza & Lennartson, 2005). Thus, it seems clear that participants in within-language conditions were more prone to distort their memory in their most proficient language (for a review, see Suarez & Beato, 2021). These results go in line with the Revised Hierarchical Model (Kroll & Stewart, 1994) that states that the higher the proficiency, the greater the automatic access to the conceptual representations, leading to higher false memory in the L1 than in the L2.

Secondly, research including both within- and between-language conditions was interested in understanding what happened with false memories when study and test language matched and did not match. This comparison between within- and between-language conditions has been called the effect of language shift. Studies that have analyzed this effect have not reached a clear conclusion, as all possible results have been reported. Specifically, it is possible to find studies with higher false recognition in within- than between-language conditions (Cabeza & Lennartson, 2005; Sahlin et al., 2005), but also studies with the opposite result (Howe et al., 2008; Marmolejo et al., 2009). It was even possible to find similar false memory for within- and between-language conditions, although this difference was not statistically tested (see Figure 1 in Kawasaki-Miyaji et al., 2003).

The effect of memory instructions: restrictive versus inclusive instructions

After carefully analyzing the previous literature on within- and between-language false memory, we found that they differ in an important respect: the instructions given to the participants at the memory test. Researchers have used two types of instructions that triggered different strategies at retrieval, and that led participants to respond based on two very different criteria. We have called these instructions *restrictive* and *inclusive* memory instructions.

In *restrictive* memory instructions, participants were asked to endorse the studied items in the memory test only when the language at study and test matched, that is, in within-language conditions (Cabeza & Lennartson, 2005; Kawasaki-Miyaji et al., 2003; Sahlin et al., 2005).[2] Therefore, when study and test languages did not match, that is, in between-language conditions, participants should reject translated studied words. Thus, with these restrictive instructions, participants are explicitly asked to make judgments requiring retrieval of language information to confirm whether the language at study and test match. Hence, with restrictive instructions, participants

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

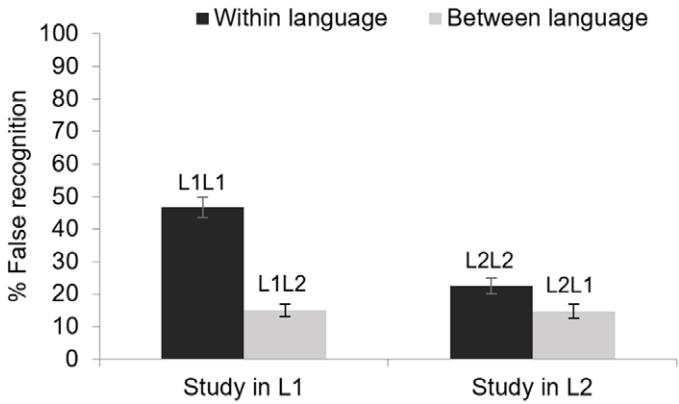

Figure 1. Percentage of false recognition in within- and between-language conditions by study language (L1: Spanish vs. L2: English) in Experiment 1.

would adopt a more conservative criterion during the recognition test by engaging in source-monitoring processes about the study language (i.e., AMF) or by searching for verbatim traces (i.e., FTT). These two mechanisms could help avoid false memories, especially in between-language conditions.

In contrast, *inclusive* memory instructions led participants to endorse all the studied items, even when presented in different languages at study and test (Howe et al., 2008). Thus, these inclusive instructions should promote a more lenient criterion at the recognition test, since participants should endorse studied words regardless of whether the study and test language matched or not. Therefore, with inclusive instructions, it was not necessary to retrieve language information to check whether the study and test languages were the same. Hence, inclusive instructions in within- and between-languages conditions give us a clearer picture of the activation of conceptual representations (i.e., AMF) or the strength of the gist traces (i.e., FTT). Taking into account our assumption that restrictive and inclusive memory instructions triggered different strategies at retrieval and led participants to respond according to two different criteria, it is not surprising that previous literature has reported mixed findings.

Regarding the effect of language shift on false memory, on the one hand, with restrictive memory instructions, we expect to find lower false recognition in between- than within-language conditions as participants should only endorse studied items when study and test language matches. Therefore, even if participants falsely recognize a critical lure as a studied item in between-language conditions, they would reject it at test when they identify it as a translated studied item (i.e., by correctly retrieving the study language). Indeed, this was the result found in previous studies that used this type of instructions (Cabeza & Lennartson, 2005; Sahlin et al., 2005).

On the other hand, with inclusive memory instructions, previous studies by Howe et al. (2008) and Marmolejo et al. (2009) have found higher false recognition in between- than within-language conditions, although these studies raise some concerns that we should be aware of. First, the false recognition obtained by

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

Howe et al. (2008) is challenging to interpret because the authors included in the between-language conditions two conditions in which study and recognition test language matched. Specifically, the L1L2L1 and L2L1L2 conditions (study-recall-recognition language) were included as between-language conditions when, in our opinion, they were rather within-language conditions in terms of the recognition memory test. Second, although Marmolejo et al. (2009) concluded that false recognition was higher in between-language conditions (i.e., L1L2 and L2L1) than in within-language conditions (i.e., L1L1 and L2L2), they did not find differences between any of the two comparisons of interest: L1L1 versus L1L2 (.80 vs. .87, respectively) and L2L2 versus L2L1 (.73 vs. .80, respectively). Third, both studies employed a free recall task before the recognition test, and therefore, the false recognition should be interpreted cautiously as these results could be contaminated by the preceding recall memory test (e.g., Gallo, 2006; Roediger & McDermott, 1995). Finally, it should be noted that these studies included participants with high L2 proficiency and, therefore, their results may not generalize to unbalanced bilinguals or second-language learners, as those included in our study.

In summary, it is not possible to get a clear picture of the effect of language shift on false memories based on previous research because the only five studies that have compared false memory in within- and between-language conditions have used different instructions and have sometimes included different memory tests (Cabeza & Lennartson, 2005; Howe et al., 2008; Kawasaki-Miyaji et al., 2003; Marmolejo et al., 2009; Sahlin et al., 2005). Therefore, the main goal of the present study was to examine the effect of language shift (within- vs. between-language conditions) on false recognition while manipulating the memory instructions. Specifically, when restrictive instructions were used (Experiment 1), participants needed to make judgments requiring retrieval of language information, whereas with inclusive instructions (Experiment 2), participants did not need to retrieve the language information. It is worth noting that only a recognition test was included in the retrieval phase to avoid any possible data contamination by an initial free recall test. To our knowledge, this is the first time in the literature that research examines the role that restrictive and inclusive memory instructions play on the effects of language shift on false recognition in second-language learners.

Experiment 1

In Experiment 1, we studied the effect of language shift on false recognition when participants should make judgments requiring retrieval of the study language. Specifically, second-language learners were presented with restrictive memory instructions that led them to endorse an item in the recognition test only when the language at study and test matched (i.e., within-language conditions) but not when it did not match (i.e., between-language conditions).

As in previous research that used within-language conditions, and in which highly proficient participants in the L1 with less proficiency in the L2 were included, we expected to find higher false recognition in L1L1 than in L2L2 (see Suarez & Beato, 2021 for a review). This outcome could be explained by the RHM (Kroll & Stewart, 1994). According to this model, the activation of the critical lures will

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

be greater when words are studied in the L1 than in the L2 due to the greater automatic access to the conceptual representations in the L1. In turn, false recognition will be higher in L1L1 than in L2L2.

More importantly, regarding the effect of language shift, with these restrictive memory instructions, we expected to find lower false recognition in between- than in within-language conditions, as in previous studies (Cabeza & Lennartson, 2005; Sahlin et al., 2005), both when words were studied in the L1 (i.e., L1L1 > L1L2) and L2 (i.e., L2L2 > L2L1). This is because even though participants might falsely recognize a critical lure in between-language conditions as a translated studied word, assuming that they are able to retrieve the study language, the instructions themselves would lead them to reject that word because it is not in the same language in which they studied it.

Method

Participants

Ninety undergraduate students (19 to 35 years old, $M = 20.61$, $SD = 2.88$) participated voluntarily and signed an informed consent form. Participants were native Spanish speakers (72% women) and were living in Spain at the time of the experiment. The Spanish school system includes mandatory English language training in primary and secondary school, but the usual situation is that young adults do not speak or listen to English on a daily basis outside of formal instruction. To measure Spanish (L1) and English (L2) proficiency, a self-report scale ranging from 1 (elementary knowledge) to 10 (native speaker proficiency) was used. All participants self-rated their L1 proficiency with a perfect score ($M = 10$, $SD = 0.00$) and their L2 proficiency as moderate,[3] giving scores that ranged from 1 to 8, with only one participant rating 10 ($M = 5.68$, $SD = 1.70$). Taking this into account, participants were considered unbalanced bilinguals, being Spanish their dominant language and English their non-dominant language. The study was approved by the Research Ethics Committee of the University of Salamanca.

Materials

Sixteen DRM lists with 10 words per list were used in the study phase (materials are freely available at https://osf.io/bz4g9/), 8 lists in Spanish and 8 lists in English (see Appendix). Spanish lists were built using the Fernandez et al. (2003) free association norms for Spanish words and normed on native Spanish speakers (Alonso et al., 2004). For their part, English lists were built using the Nelson et al. (1998) free association norms for English words and normed on native English speakers (Stadler et al., 1999).

Since we were aware of the traditionally high variability in false memory rates produced by DRM lists, we decided to match our L1 and L2 DRM lists for their level of false recognition. That is, Spanish and English lists had similar false recognition rates (54.00% and 54.38%, respectively) when they were applied to native-speaker participants in previous normative studies (Alonso et al., 2004; Stadler et al., 1999), $t(14) = -.048$, $p = .962$, Cohen's $d = 0.03$, 95% CI [−0.17, 0.16]. In other words, Spanish and English lists share the same capacity to produce false

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

recognition. Consequently, if the lists show different levels of false recognition in the present study, it is unlikely that these differences are due to the lists. Furthermore, eight DRM lists (four in English and four in Spanish) with similar false recognition rates in both languages were used as unrelated distractors in the recognition memory test.

The 96-item recognition memory test included 48 studied words (three per study list, serial positions 1, 6, and 10) and 48 non-studied words (16 critical lures and 32 unrelated distractors), half in English and half in Spanish. Words were randomly presented. Half of the critical lures were presented in the same language as their corresponding studied lists (i.e., within-language conditions: L1L1 or Spanish-Spanish, and L2L2 or English-English). The other half of the critical lures was translated into the other language (i.e., between-language conditions: L1L2 or Spanish-English, and L2L1 or English-Spanish). Critical lures were evenly assigned to within- and between-language conditions.

Procedure

Participants were run in groups of up to 22 and were presented with 16 study lists (eight in Spanish and eight in English). Each word was visually presented on a computer screen for 2 s. The associates within each list were presented in decreasing order of associative strength, and lists were randomly presented. Participants were instructed in Spanish (L1) to pay attention to each word in preparation for a subsequent memory test (approximated English translation): "In this part of the experiment, words will be presented one at a time on the computer screen. Words will appear at a constant rate in the center of the screen, pay attention to them. Your task is to study the words as best you can because afterward you will have to perform a memory test. Some words will be presented in English and others in Spanish. Do you have any questions?"

Following the study phase, participants were administered the recognition test, in which words were presented one at a time on the computer screen. Participants were asked to decide whether each word had been presented (i.e., "old" word) or not (i.e., "new" word) at the study phase and to respond by pressing the corresponding key. As in previous research (Cabeza & Lennartson, 2005), restrictive memory instructions were provided, and participants had to respond "old" only to the studied words that appeared in the test phase in the same language as in the study phase. Otherwise, participants would have to respond "new" to the word. The approximated English translation of the instructions would be: "In this part of the experiment, we will test your memory. You will be presented with words one at a time on the computer screen. Your task is to determine whether each word was presented, or not, in the study phase. If the word was presented in the list of words that you just studied, please press the "E" key to indicate that it was STUDIED. If the word was not presented in the list of words you just studied, please press the "N" key to indicate that it is NEW. Be careful because you may have studied a word in Spanish and now, in the memory test, the word appears translated into English. In this case, you must indicate that it is a NEW word and press the "N" key. That is, you have to consider as STUDIED only the words that are written in the same language in which they were previously studied. For example, if you studied the word "FRANCIA" and

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

Table 1. Mean percentage of false recognition with restrictive memory instructions (Experiment 1) and inclusive memory instructions (Experiment 2) as a function of study and test language

	Study and test language condition			
Memory instructions	L1L1	L2L2	L1L2	L2L1
Restrictive instructions	46.67 (29.79)	22.50 (24.01)	15.00 (18.66)	14.72 (19.44)
Inclusive instructions	56.94 (29.28)	34.44 (25.58)	47.50 (26.25)	38.89 (29.06)

Note. L1 = Spanish; L2 = English. Standard deviations are reported in parenthesis.

the word "FRANCE" appears in the memory test, you should consider it as a NEW word by pressing the "N" key. You must do the same with the words that you studied in English and now are presented in Spanish. Do you have any questions?"

Results and discussion

Complete analysis code and data are freely available at https://osf.io/bz4g9/

Language dominance effect on true memory

To analyze language dominance effect on true memory, we compared the percentages of "old" responses to studied words in a dominant language or L1 (i.e., Spanish) and a non-dominant language or L2 (i.e., English). The paired t-test indicated that participants correctly recognized fewer words studied in the L1 ($M = 68.84$, $SD = 16.95$) than in the L2 ($M = 77.45$, $SD = 13.66$), $t(89) = -5.484$, $p < .001$, Cohen's $d = -0.56$, 95% CI [−11.73, −5.49]. Previous studies conducted with second-language learners have reported a similar pattern of results, that is, higher true recognition in the non-dominant language (e.g., Arndt & Beato, 2017; Beato & Arndt, 2021).

Language dominance effect on false memory

To evaluate whether there was a language dominance effect on false recognition in both within- and between-language conditions, we conducted two separate two-way ANOVAs (see Table 1 for descriptives).

First, a 2 (study language: L1 [Spanish], L2 [English]) x 2 (type of word: critical lure, unrelated distractor) repeated-measures ANOVA was performed on the percentage of "old" responses given to each type of word in within-language conditions. The analysis revealed a significant main effect of study language, $F(1, 89) = 34.38$, $p < .001$, $\eta^2_p = .279$, with more "old" responses in the L1 than in the L2 condition (26.20 vs. 16.44, respectively), 95% CI [6.46, 13.08]. A significant main effect of type of word, $F(1, 89) = 122.85$, $p < .001$, $\eta^2_p = .580$, showed that "old" responses to critical lures ($M = 34.58$) were more likely than "old" responses to unrelated distractors ($M = 8.06$), 95% CI [21.77, 31.28]. Finally, there was a significant Study Language × Type of Word interaction, $F(1, 89) = 86.30$, $p < .001$, $\eta^2_p = .492$. False alarms to critical lures were higher than false alarms to unrelated distractors in both the L1 (46.67 vs. 5.74, respectively), 95% CI [34.77, 47.08], $p < .001$, Cohen's $d = 1.87$, and the L2 (22.50 vs. 10.37, respectively), 95% CI [7.00, 17.26], $p < .001$,

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

Cohen's $d = 0.65$, but the interaction occurred because the effect of type of word was larger in the L1 than in the L2 condition. These data confirmed that critical lures, regardless of the language, produced above-baseline levels of false recognition. Moreover, as expected, false recognition was higher in critical lures associated with words studied in the L1 (i.e., L1L1 condition, $M = 46.67$) than in the L2 condition (i.e., L2L2 condition, $M = 22.50$), $p < .001$, Cohen's $d = 0.89$, 95% CI [18.12, 30.22]. False alarms to unrelated distractors were higher in the L2 ($M = 10.37$) than in the L1 ($M = 5.74$), $p < .001$, Cohen's $d = 0.47$, 95% CI [2.57, 6.69], a trend opposite to that for critical lures.

Second, we analyzed the language dominance effect on false recognition in between-language conditions. In these conditions, the critical lures at test were presented in a different language to that in which their associates were studied, and participants should give "old" responses only to those words that matched language at study and test. The 2 (study language: L1, L2) × 2 (type of word: critical lure, unrelated distractor) repeated-measures ANOVA performed on the percentage of "old" responses given to each type of word in between-language conditions revealed a significant main effect of type of word, $F(1, 89) = 18.41$, $p < .001$, $\eta^2_p = .171$, where false alarms to critical lures ($M = 14.86$) were higher than false alarms to unrelated distractors ($M = 8.06$), 95% CI [3.65, 9.96]. Therefore, we confirmed that, as in previous studies with the same restrictive memory instructions, there was false recognition in between-language conditions (e.g., Cabeza & Lennartson, 2005). In contrast, neither the main effect of study language, $F(1, 89) = 2.66$, $p = .106$, $\eta^2_p = .029$ nor the Study Language × Type of Word interaction, $F(1, 89) = 3.41$, $p = .068$, $\eta^2_p = .037$ were statistically significant.

In summary, using restrictive memory instructions that required participants to retrieve the study language, critical lures were falsely recognized in both within- and between-language conditions, showing a false memory effect. It is worth mentioning that in the between-language conditions, not only critical lures were not presented at study but also appeared translated in the recognition test and, despite all this, false memory was found, confirming the robustness of this effect. This finding shows that participants were not able to retrieve the language information of all the studied lists, as they were unable to reject all the translated critical lures. Furthermore, a language dominance effect was only found in within-language conditions. Specifically, false recognition was higher when words were studied in the dominant than in the non-dominant language as long as the study and test language matched (e.g., Anastasi, Rhodes, et al., 2005; Arndt & Beato, 2017; Beato & Arndt, 2021; Sahlin et al., 2005), but there were no differences when the languages at study and test did not match (e.g., Kawasaki-Miyaji et al., 2003; Sahlin et al., 2005).

Effect of language shift on false memory

In order to examine the effect of language shift on false memory, we conducted a 2 (study language: L1, L2) × 2 (language shift: within-language conditions [L1L1, L2L2], between-language conditions [L1L2, L2L1]) ANOVA on false recognition (i.e., false alarms to critical lures). This ANOVA yielded a significant main effect

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

of study language, $F(1, 89) = 47.28$, $p < .001$, $\eta^2_p = .347$, where false recognition was higher when words were studied in the L1 ($M = 30.83$) than in the L2 ($M = 18.61$), 95% CI [8.69, 15.75]. Also, there was a main effect of language shift, $F(1, 89) = 56.02$, $p < .001$, $\eta^2_p = .386$, showing a higher false recognition in within-language conditions (i.e., L1L1 and L2L2) than in between-language conditions (i.e., L1L2 and L2L1) (34.58 vs. 14.86, respectively), 95% CI [14.49, 24.96].

Finally, there was a significant Study Language × Language Shift interaction, $F(1, 89) = 31.82$, $p < .001$, $\eta^2_p = .263$. As can be seen in Figure 1, false recognition was higher when words were studied in the L1 ($M = 46.67$) than in the L2 ($M = 22.50$) in within-language conditions, $p < .001$, Cohen's $d = 0.90$, 95% CI [18.12, 30.22], but not in between-language conditions (15.00 vs. 14.72, respectively), $p = .910$, Cohen's $d = 0.02$, 95% CI [−4.59, 5.15]. Furthermore, regarding the effect of language shift, false recognition was higher in within- than in between-language conditions when words were studied in both the L1 (L1L1: 46.67 vs. L1L2: 15.00), $p < .001$, Cohen's $d = 1.28$, 95% CI [24.61, 38.72], and the L2 (L2L2: 22.50 vs. L2L1: 14.72), $p = .017$, Cohen's $d = 0.36$, 95% CI [1.42, 14.14], but this difference was higher when words were studied in the L1.

Therefore, regarding the effect of language shift when restrictive memory instructions were used, the results indicated that false recognition was higher in within- than in between-language conditions regardless of the language in which the words were studied, replicating the findings of previous research that used the same restrictive instructions (Cabeza & Lennartson, 2005; Sahlin et al., 2005). There are two possible explanations for this result: there was less false memory formation in between- than in within-language conditions or there was a correct retrieval of language information in between-language conditions. One could argue that participants rejected a higher number of critical lures in between-language conditions due to a lack of memory traces for those concepts. However, when we compare within- and between-language conditions (i.e., L1L1 vs. L1L2, or L2L2 vs. L2L1), we are contrasting conditions in which the study phase was the same. Since activation and gist extraction mainly occur in the study phase, in the comparison between L1L1 versus L1L2 (or L2L2 vs. L2L1), we expect that the critical lures were equally activated or the gist memory traces were the same in both conditions. If that was the case, we suggest that participants rejected more critical lures in between- than in within-language conditions because the restrictive memory instructions forced them to do so by correctly retrieving the study language, the second possible explanation. That is, it could be that participants (falsely) thought that critical lures were studied items but presented in a different language to that used at study and as the language did not match at study and test, following the instructions, participants rejected those critical lures. In this sense, the restrictive memory instructions made it difficult to know what was happening with critical lures: In between-language conditions, did participants have memory traces for the rejected critical lures, or did they not? Experiment 2 sought to answer this question by analyzing false recognition in within- and between-language conditions with different memory instructions that will allow us to fully capture the false memory illusion.

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

Experiment 2

Experiment 2 studied the effect of language shift on false recognition when second-language learners do not need to retrieve the study language. Concretely, participants were presented with inclusive memory instructions that consisted of responding "old" in the recognition test when the item had been presented in the study phase, regardless of whether the study and test language matched or not. With these inclusive instructions, first, we expected to find, as in Experiment 1, higher false recognition when words were studied in the L1 (i.e., L1L1) than in the L2 (i.e., L2L2). Second, regarding the effect of language shift on false recognition, we consider that, in Experiment 1, participants had memory traces for the rejected critical lures in between-language conditions, but they were rejected because participants correctly retrieved the study language. In contrast, in Experiment 2, as participants have to endorse all studied words regardless of the language in which they were presented (i.e., inclusive memory instructions), we expect to find an increase in between-language false recognition (i.e., L1L2 and L2L1) reaching a similar level to the false recognition in within-language conditions (i.e., L1L1 and L2L2).

Method

Participants

We recruited 90 undergraduate students (85.56% women), native Spanish speakers living in Spain at the time of the experiment. None of them had participated in Experiment 1. Participants' ages ranged from 19 to 29 years ($M = 20.24$; $SD = 2.21$). They self-rated their proficiency (using a 10-point scale) in Spanish (L1) with a perfect score ($M = 10$, $SD = 0.00$) and their proficiency in English (L2) close to the midpoint ($M = 4.94$, $SD = 1.65$), with scores ranging from 1 to 8. All participants were volunteers and signed an informed consent form. The study was approved by the Research Ethics Committee of the University of Salamanca.

Materials

In Experiment 2, we used the same materials as in Experiment 1.

Procedure

The procedure was similar to that used in Experiment 1, except for the instructions provided to the participants in the recognition test. In this experiment, participants were instructed in their L1 to respond "old" to previously presented words in the study phase, regardless of the language in which those words had appeared (i.e., inclusive memory instructions). That is, they should endorse the items presented at the study phase, regardless of whether the study and test language matched or not. Similar instructions have also been used in Howe et al.'s (2008) study. The approximated English translation of the instructions for Experiment 2 would be: "In this part of the experiment, we will test your memory. You will be presented with words one at a time on the computer screen. Your task is to determine whether each word was presented, or not, in the study phase. If the word was presented in the

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

list of words that you just studied, please press the "E" key to indicate it is STUDIED. If the word was not presented on the list of words you just studied, please press the "N" key to indicate it is NEW. Be careful because you may have studied a word in Spanish and now in the memory test the word appears translated into English. In this case, you must indicate that it was STUDIED and press the "E" key. That is, you have to consider as STUDIED the words that were presented in the study phase, regardless of the language (Spanish or English). For example, if you studied the word "FRANCIA" and the word "FRANCE" appears in the memory test, you should consider it as a STUDIED word by pressing the "E" key. You must do the same with the words that you studied in English and now are presented in Spanish. Do you have any questions?"

Results and discussion

Complete analysis code and data are freely available at https://osf.io/bz4g9/

Language dominance effect on true memory

In order to check whether, as in Experiment 1, true memory was significantly higher in the non-dominant language than in the dominant language, we compared the "old" responses to words studied in the L1 (i.e., Spanish) and L2 (i.e., English). The paired t-test indicated that true recognition was higher in words studied in the L2 ($M = 80.93$, $SD = 12.43$) than in the L1 ($M = 74.77$, $SD = 15.47$), $t(89) = -3.45$, $p = .001$, Cohen's $d = -0.44$, 95% CI [−9.70, −2.62]. This result replicated findings from our Experiment 1 and previous studies conducted with second-language learners (e.g., Arndt & Beato, 2017; Beato & Arndt, 2021).

Language dominance effect on false memory

As in Experiment 1, to assess whether there was a language dominance effect on false recognition in both within- and between-language conditions, we conducted two separate two-way ANOVAs (see Table 1).

First, a 2 (study language: L1, L2) × 2 (type of word: critical lure, unrelated distractor) repeated-measures ANOVA was performed on the percentage of "old" responses in within-language conditions. This analysis revealed a significant main effect of study language, $F(1, 89) = 24.73$, $p < .001$, $\eta^2_p = .217$, showing a higher percentage of "old" responses in the L1 than in the L2 condition (33.10 vs. 24.44, respectively), $p < .001$, 95% CI [5.20, 12.12]. A significant main effect of type of word, $F(1, 89) = 195.03$, $p < .001$, $\eta^2_p = .687$, showed that "old" responses to critical lures ($M = 45.69$) were more likely than "old" responses to unrelated distractors ($M = 11.85$), 95% CI [29.03, 38.66], confirming the existence of false recognition. Finally, there was a significant Study Language × Type of Word interaction, $F(1, 89) = 57.50$, $p < .001$, $\eta^2_p = .392$. Specifically, as in Experiment 1, this interaction indicated that although "old" responses to critical lures were higher than "old" responses to unrelated distractors in both the L1 (56.94 vs. 9.26, respectively), $p < .001$, Cohen's $d = 2.19$, 95% CI [41.43, 53.94], and L2 (34.44 vs. 14.44, respectively), $p < .001$, Cohen's $d = 0.97$, 95% CI [14.20, 25.80], this difference was greater in the L1 than in the L2. Moreover, as in Experiment 1, false recognition

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

was higher in critical lures associated with words studied in the L1 ($M = 56.94$) than in the L2 ($M = 34.44$), $p < .001$, Cohen's $d = 0.82$, 95% CI [15.88, 29.12]. Also as in Experiment 1, false alarms to unrelated distractors had an inverse pattern with higher values in the L2 ($M = 14.44$) than in the L1 condition ($M = 9.26$), $p < .001$, Cohen's $d = 0.43$, 95% CI [2.64, 7.73].

Therefore, Experiment 2 replicated the results of Experiment 1 in the within-language conditions. That is, we found false recognition in both dominant and non-dominant languages. Furthermore, false recognition proved to be higher in the dominant than in the non-dominant language (e.g., Howe et al., 2008).

Second, regarding the language dominance effect on false recognition in between-language conditions, a 2 (study language: L1, L2) × 2 (type of word: critical lure, unrelated distractor) repeated-measures ANOVA was performed on the percentage of "old" responses. As in Experiment 1, the ANOVA revealed a significant main effect of type of word, $F(1, 89) = 244.94$, $p < .001$, $\eta^2_p = .733$, with higher "old" responses to critical lures ($M = 43.19$) than to unrelated distractors ($M = 11.85$), $p < .001$, 95% CI [27.36, 35.32]. In addition, there was no significant main effect of study language, $F(1, 89) = 0.79$, $p = .378$, $\eta^2_p = .009$. Finally, and in contrast to Experiment 1, there was a significant Study Language × Type of Word interaction, $F(1, 89) = 12.01$, $p = .001$, $\eta^2_p = .119$. Specifically, Bonferroni post hoc tests indicated that there was a false memory effect in the dominant and the non-dominant language, with "old" responses to critical lures higher than "old" responses to unrelated distractors in both the L1 (47.50 vs. 9.26, respectively), $p < .001$, Cohen's $d = 1.94$, 95% CI [33.00, 43.48], and L2 (38.89 vs. 14.44, respectively), $p < .001$, Cohen's $d = 1.07$, 95% CI [18.49, 30.41]. Again, the difference between critical lures and unrelated distractors was larger in the L1 than in the L2. Furthermore, analyses revealed that false recognition was higher in critical lures associated with words studied in the L1 ($M = 47.50$) than in the L2 ($M = 38.89$), $p = .023$, Cohen's $d = 0.31$, 95% CI [1.24, 15.98], but false alarms to unrelated distractors were higher in the L2 ($M = 14.44$) than in the L1 ($M = 9.26$), $p < .001$, Cohen's $d = 0.43$, 95% CI [2.64, 7.73]. Thus, unlike what happened in the same condition in Experiment 1, we also found a language dominance effect on false recognition in the between-language conditions. That is, the use of inclusive memory instructions allowed us to find that critical lures associated with words studied in the L1 were more falsely recognized than critical lures whose associates were studied in the L2, not only in within-language conditions but also in between-language conditions.

Effect of language shift on false memory

To analyze the effect of language shift on false recognition, a 2 (study language: L1, L2) × 2 (language shift: within-language conditions [L1L1, L2L2], between-language conditions [L1L2, L2L1]) ANOVA was carried out. The analysis revealed a significant main effect of study language, $F(1, 89) = 34.78$, $p < .001$, $\eta^2_p = .281$. Specifically, as in Experiment 1, false recognition was higher when words were studied in the L1 ($M = 52.22$) than in the L2 ($M = 36.67$), 95% CI [10.32, 20.80]. Moreover, as expected, there was no significant main effect of language shift, $F(1, 89) = 0.808$, $p = .371$, $\eta^2_p = .009$, 95% CI [−3.03, 8.03].

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

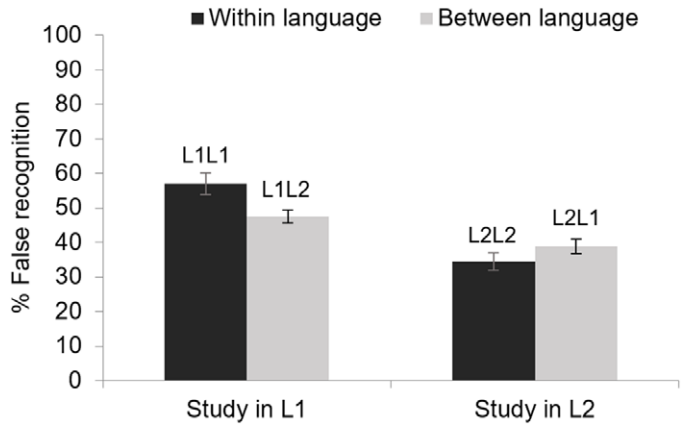

Figure 2. Percentage of false recognition in the within- and between-language conditions by study language (L1: Spanish vs. L2: English) in Experiment 2.

Finally, there was a significant Study Language × Language Shift interaction, $F(1, 89) = 8.82$, $p = .004$, $\eta^2_p = .09$. Bonferroni post hoc tests showed that false recognition was higher when words were studied in the L1 than in the L2, both in within-language conditions (i.e., L1L1 vs. L2L2) (56.94 vs. 34.44, respectively), $p < .001$, Cohen's $d = 0.82$, 95% CI [15.88, 29.12], and in between-language conditions (i.e., L1L2 vs. L2L1) (47.50 vs. 38.89), $p = .023$, Cohen's $d = 0.31$, 95% CI [1.24, 15.98], but this difference was much larger for within- than between-language conditions (see Figure 2). Furthermore, regarding the effect of language shift, false recognition was higher in within- than in between-language conditions only when words were studied in the L1 (i.e., L1L1 vs. L1L2; 56.94 vs. 47.50, respectively), $p = .016$, Cohen's $d = 0.34$, 95% CI [1.77, 17.12], but not in the L2 (i.e., L2L2 vs. L2L1; 34.44 vs. 38.89, respectively), $p = .193$, Cohen's $d = -0.16$, 95% CI [−11.18, 2.29].

General discussion

The aim of this research was to examine the effect of language shift on false recognition (i.e., to compare false recognition in conditions in which study and test language matched or not) in second-language learners while manipulating the memory instructions, so participants needed or not to retrieve language information to make the memory judgments.

For this purpose, we conducted two experiments in which we manipulated the instructions given at the recognition test, while the study phase instructions always remained the same. Specifically, in Experiment 1, we used restrictive memory instructions that required participants to endorse the studied items only when the language at study and test was the same (i.e., within-language conditions: L1L1, L2L2) and therefore, to reject the translated studied items (i.e., between-language conditions: L1L2, L2L1). In other words, with restrictive memory instructions, participants were forced to make judgments requiring retrieval of language

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

information (i.e., to engage in source-monitoring processes, as predicted by the AMF, or to search for verbatim traces, as predicted by the FTT). In Experiment 2, inclusive memory instructions were used. These instructions required participants to endorse all the studied items regardless of the language in which they were presented. Thus, with inclusive memory instructions, participants were not required to make judgments requiring retrieval of word-specific language information (whether that be via monitoring or retrieval of verbatim traces), as they had to endorse the studied items in both within- and between-language conditions.

Regarding true memory, in both experiments, we found that true recognition was higher in the L2 or non-dominant language than in the L1 or dominant language, replicating previous findings of studies conducted with second-language learners or unbalanced bilinguals (e.g., Arndt & Beato, 2017; Beato & Arndt, 2021). Specifically, these authors suggested it is more effortful to access concepts' lexical and semantic representations in the non-dominant than in the dominant language. This higher difficulty level benefits true recognition in the L2 as compared to the L1 (e.g., Bjork & Bjork, 2011).

Looking into the false memory effect, we found false recognition in within- and between-language conditions using not only inclusive, but also restrictive memory instructions. In other words, although the restrictive instructions at test promote the engagement of source-monitoring processes or the search for verbatim traces that reduce the false memory effect, this effect appeared with those restrictive instructions, even in the between-languages conditions. It may seem surprising that participants falsely recognized critical lures in between-language conditions when they were instructed to reject translated studied words (i.e., restrictive memory instructions). This finding can be explained by the AMF and the FTT. On the one hand, the AMF states that, in the DRM paradigm, memory distortions arise when the activation of preexisting associations between the studied words and the critical lure occurs (activation processes that are automatic to some extent), and subsequently, monitoring processes fail. On the other hand, the FTT posits that memory distortions appear when gist information of the list is extracted, this gist trace matches the critical lure, and the retrieval of verbatim traces of the studied words is not enough to reject the critical lure. Furthermore, our false memory results with restrictive memory instructions seem to indicate that participants were able to engage in some language source monitoring (AMF), or to retrieve verbatim traces (FTT), as they produced a lower false memory rate when study and test language did not match (between-language conditions) than when they matched (within-language conditions). However, the false memory effect with the DRM paradigm is so robust that these instructions could not eliminate the effect (see previous research using warning instructions for similar results; e.g., Carneiro & Fernandez, 2010; McDermott & Roediger, 1998; Watson et al., 2004).

We were also interested in examining the language dominance effect on false memory for within- and between-language conditions. On the one hand, in within-language conditions, as expected, we found a higher false recognition in the dominant (i.e., L1L1) than in the non-dominant language (i.e., L2L2) regardless of whether participants had or did not have to retrieve the language information (restrictive and inclusive memory instructions, respectively). In other words, a language dominance effect was found in false recognition in within-language

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

conditions, as in previous studies (e.g., Anastasi, Rhodes, et al., 2005; Arndt & Beato, 2017; Sahlin et al., 2005; for a review, see Suarez & Beato, 2021), using restrictive as well as inclusive memory instructions. On the other hand, in between-language conditions, we also observed a higher false recognition for critical lures associated with words studied in the dominant language (i.e., L1L2) than in the non-dominant language (i.e., L2L1), but only when participants were required to endorse the studied items regardless of the language in which they were presented. That is, we found a language dominance effect in false recognition in between-language conditions only when inclusive memory instructions (Experiment 2) were used.

The findings regarding the language dominance effect on false memory could be explained in terms of the RHM (Kroll & Stewart, 1994), the AMF (Roediger et al., 2001), and the FTT (Brainerd & Reyna, 2002). In relation to the RHM and AMF, conceptual representations will be activated more quickly and automatically from L1 than L2 words. Thus, the conceptual representation of the critical lures would be more activated if the lists were studied in the L1 than in the L2. For its part, according to the FTT, it can be assumed that there is a greater strength of gist memory traces in L1 than in L2 word lists because gist traces are supposed to store many conceptually based elements of an experience, such as its meaning (Arndt & Gould, 2006). These arguments would explain the language dominance effect on false memory in within-language conditions. In the between-language conditions with restrictive instructions, participants had to retrieve the study language and, since study and test languages did not match, they had to reject the falsely recognized translated critical lures. This would explain why no language dominance effect on false memory was found in between-language conditions with restrictive instructions. However, this effect was found in between-language conditions when using inclusive instructions since participants did not have to engage in language source-monitoring processes or search for verbatim traces. That is, inclusive instructions allowed us to capture either all the activation of critical lures or the strength of gist memory traces, being in both cases stronger in the L1 than the L2.

The difference in the language dominance effect in between-language conditions when using restrictive versus inclusive memory instructions (Experiment 1 and 2, respectively) provides evidence of the existence of episodic details in false memories. To understand how this result pattern shows evidence for recollection processes in false memories, we present the following example. In between-language conditions, participants studied a list, for example, in their L1 (mesa, sillón, sentarse, descanso, asiento, cansancio, sofá, taburete, cocina, respaldo) and were presented in the recognition test with the critical lure in their L2 (e.g., CHAIR). First, with inclusive instructions, the L1L2 condition showed higher levels of false recognition than the L2L1 condition (i.e., language dominance effect). That is, the critical lure CHAIR (L2), when its list was studied in the L1, would be more endorsed than, for example, the critical lure AGUJA (L1) when its list was studied in the L2 (thread, pin, eye, sewing, sharp, point, prick, thimble, haystack, thorn). This result can be explained in terms of the AMF and the FTT. According to the AMF, the conceptual representations of critical lures derived from studied lists in the dominant language (L1L2) would receive a more pronounced activation than in the non-dominant language (L2L1). Regarding the FTT, these data show that the gist memory traces from studied lists in the dominant language (L1L2) were stronger than those from studied

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

lists in the non-dominant language (L2L1). Second, the fact that the language dominance effect was not found when using restrictive instructions (Experiment 1), as noted above, seems to be due to the fact that participants were able to engage in language source monitoring (AMF) or to retrieve verbatim memory traces (FTT). That is, in L1L2 and L2L1 conditions, when provided with restrictive instructions, participants seem to avoid committing false memories to some extent by recollecting the language in which the critical lure was "studied" (the same language in which the associates of that list were studied). In other words, participants seem to retrieve language information as episodic details (i.e., source-monitoring in AMF or verbatim traces in the FTT) about their false memory to avoid identifying some critical lures as studied items in between-language conditions.

Lastly, the results of the effect of language shift on false recognition will be discussed. In Experiment 1 with restrictive instructions, we found higher false recognition in within- than in between-language conditions, regardless of the language in which the words were studied (i.e., L1L1 > L1L2, and L2L2 > L2L1). This finding replicated the results of previous studies that have used these restrictive instructions (Cabeza & Lennartson, 2005; Sahlin et al., 2005), and there are two possible explanations for it. First, it could be that in between-language conditions participants were quite resistant to the DRM false memory illusion because study and test languages did not match. Second, it could be the case that participants rejected more critical lures in between- than within-language conditions because they were able to retrieve the language information and (falsely) thought that critical lures were translated studied items. In other words, participants had false memories of those critical lures but answered "no" in the recognition test because they could remember the language in which the critical lures' lists were presented and did not match the test language. The difference between these two possible explanations is substantial since there would be no false recognition in the former, while in the latter, there would be.

To disentangle whether or not there was false recognition, in Experiment 2, we used inclusive memory instructions. These instructions did not require participants to retrieve the study language and would allow them to endorse translated critical lures in case they have (false) memory for those critical lures. Therefore, when there is no need to retrieve the language, we can capture the false memory illusion more accurately. As expected, with inclusive memory instructions, in general, we found an increase in false recognition rates in between-language conditions (i.e., L1L2, L2L1) up to a similar level to the false recognition in within-language conditions (i.e., L1L1, L2L2). This increase in false recognition in between-language conditions when there was no need to retrieve the language (inclusive memory instructions) confirmed that the conceptual representations associated with critical lures in these conditions were activated, or that the gist trace was extracted and matched the critical lure. Consequently, in Expertiment 1, when the instructions did not allow participants to endorse translated studied items (restrictive instructions), they rejected more critical lures in between- than in within-language conditions because they were able to retrieve the language information (to correctly monitor the study language, according to the AMF, or to retrieve verbatim traces, following the FTT) and not because of a lack of false memories. This finding is particularly interesting as it evidences that: (1) participants created a memory trace for the critical lures, and

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

(2) the memories for the critical lures contained episodic details (i.e., the study language of the critical lures' lists).

Interestingly, analyzing the interaction between study language and language shift with inclusive instructions, we found an effect of language shift on false recognition only when words were studied in the L1 (i.e., higher false recognition in L1L1 than in L1L2), but not in the L2 (i.e., no differences in false recognition between L2L2 and L2L1). Predictions from the RHM could explain these results. According to the RHM, first, speakers usually have strong direct links from L1 words to their conceptual representations and, second, they access the conceptual representations from L2 words differently depending on their L2 proficiency. Specifically, participants with a high L2 proficiency would access the concepts directly from L2 words, while second-language learners would take a different route through the L1 translation. Thus, when participants studied words in their L1 (i.e., L1L1 and L1L2), they rapidly and automatically accessed the concepts. Furthermore, according to the AMF, that activation spread to associated concepts reaching the critical lures, or regarding the FTT, a strong gist memory trace for each list would be extracted. Later, on the one hand, when the critical lures were presented in their L1 at the recognition test (i.e., L1L1), participants falsely recognized them as studied items. On the other hand, when our participants with low L2 proficiency encountered the critical lures translated into their L2 at the recognition test (i.e., L1L2), they needed to translate those words into their L1 to access the conceptual representations. In this case, if participants were not able to translate some of the L2 words they found at test, they would not access those conceptual representations, and therefore, those translated critical lures would be rejected, finding lower false recognition in L1L2 than in L1L1 (i.e., the effect of language shift).

By contrast, when second-language learners studied words in their L2 (i.e., L2L2 and L2L1), as noted above, they accessed the conceptual representations by translating those words into their L1. Again, those conceptual representations would only be accessed when participants knew the translation of those L2 words, but in this case, the L2 knowledge would similarly affect both the L2L2 and L2L1 conditions. If the concepts associated with the critical lure were not activated (AMF) or gist traces from studied lists were not extracted (FTT) during the study phase, it does not matter in which language the critical lure is presented in the recognition test, as a similar false recognition in L2L2 and L2L1 would be expected, just as we found in our study.

This finding differs from previous studies that have used inclusive memory instructions and have found higher false recognition rates in between- than in within-language conditions (Howe et al., 2008; Marmolejo et al., 2009). This different pattern of results could be due to the fact that other researchers included bilingual participants while we tested second-language learners with a lower L2 proficiency. Moreover, as mentioned in the introduction, previous studies raised some methodological concerns that might also influence their results. Concretely, they employed a free recall task before the recognition test, and the language of this free recall task may or may not match the language of the study phase and the recognition test. Moreover, Howe et al. (2008) included as between-language conditions some conditions where the language at study and at the recognition test was the same (i.e., within-language conditions in terms of the recognition memory test).

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

Finally, although Marmolejo et al. (2009) concluded that false recognition was higher in between- than in within-language conditions, they failed to find differences between our two comparisons of interest (i.e., L1L1 vs. L1L2 and L2L2 vs. L2L1).

We would like to point out that, even though we have no translation data to prove that participants might not be correctly translating all L2 words, previous studies with Spanish second-language learners of English have shown that participants had far from perfect knowledge of all L2 words presented in the experiment (Beato & Arndt, 2021). Future research might benefit from conducting a translation test at the end of the experiment to be certain that the explanation of our results provided in this work is accurate. Additionally, further research could not only manipulate the need to make judgments requiring retrieval of word-specific language information at the test phase (as we have done in this research), but also during the study phase. The effect of language shift on false recognition is expected to be greater if, before the study phase, participants were asked to remember the language of the studied words.

In summary, we must consider two important factors when studying the effect of language shift on false memory with second-language learners. First, we need to consider in which direction the shift occurs as, according to the RHM, second-language learners use two different routes when accessing the conceptual representations from L1 and L2 words. Second, we need to bear in mind the memory instructions used because, according to our results, only the inclusive memory instructions allowed us to fully capture the false memory illusion in both within- and between-language conditions. Furthermore, only the restrictive memory instructions allowed us to provide evidence regarding the existence of episodic details in false memories.

Taken together, the current findings help us gain knowledge on the effect of language shift on false memories. The false memory effect raised with the DRM paradigm is so robust that it appears in between-language conditions tested in moderate-proficient L2 speakers. Besides, the only five studies that had compared false memory in within- and between-language conditions had reported mixed findings as they had used different memory instructions. With the present study, we get a clearer picture of the role that instructions play in producing false recognition in both within- and between-language conditions.

Replication package. All research materials, data, and analysis code are available at https://osf.io/bz4g9/

Financial support. This study was partially conducted at the Psychology Research Centre (CIPsi/UM) School of Psychology, University of Minho, supported by the Foundation for Science and Technology (FCT) through the Portuguese State Budget (UIDB/01662/2020), and partially supported by the University of Salamanca and Banco Santander through a pre-doctoral research contract attributed to MS.

Conflicts of interest. The authors declare none.

Notes

1 See also Arndt and Hirshman (1998) for a Global-Matching Model.

2 The procedure employed in Kawasaki-Miyaji et al.'s (2003) study makes it difficult to compare with other research. Specifically, a four alternative forced choice (4-AFC) recognition test was included: a) presented in English, b) presented in Japanese, c) presented but uncertain about which language, d) not presented.

3 Coinciding with the Education First (EF) English Proficiency Index: www.ef.com.

https://doi.org/10.1017/S0142716423000140 Published online by Cambridge University Press

References

Abe, N., Okuda, J., Suzuki, M., Sasaki, H., Matsuda, T., Mori, E., Tsukada, M., & Fujii, T. (2008). Neural correlates of true memory, false memory, and deception. *Cerebral Cortex*, **18**(12), 2811–2819. doi: 10.1093/cercor/bhn037

Albuquerque, P. B. (2005). Produção de evocações e reconhecimentos falsos em 100 listas de palavras associadas portuguesas [False recall and recognition production in 100 Portuguese associate word list]. *Laboratório de Psicologia*, **3**, 3–12. doi: 10.14417/lp.766

Albuquerque, P. B., & Pimentel, E. (2005). Impacto da inibição do efeito de recência na produção de memõrias falsas em listas de associados [Impact of the recency effect on false memory production in associate lists]. *Psicologia, Educação e Cultura*, **IX**, 71–90.

Alonso, M. A., Fernandez, A., Diez, E., & Beato, M. S. (2004). Índices de producción de falso recuerdo y falso reconocimiento para 55 listas de palabras en castellano [False recall and recognition indexes for 55 lists of words in Spanish]. *Psicothema*, **16**(3), 357–362. http://www.psicothema.com/pdf/3002.pdf

Anaki, D., Faran, Y., Ben-Shalom, D., & Henik, A. (2005). The false memory and the mirror effects: The role of familiarity and backward association in creating false recollections. *Journal of Memory and Language*, **52**(1), 87–102. doi: 10.1016/j.jml.2004.08.002

Anastasi, J. S., de Leon, A., & Rhodes, M. G. (2005). Normative data for semantically associated Spanish word lists that create false memories. *Behavior Research Methods*, **37**(4), 631–637. doi: 10.3758/BF03192733

Anastasi, J. S., Rhodes, M. G., Marquez, S., & Velino, V. (2005). The incidence of false memories in native and non-native speakers. *Memory*, **13**(8), 815–828. doi: 10.1080/09658210444000421

Arndt, J. (2012). The influence of forward and backward associative strength on false recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **38**(3), 747–756. doi: 10.1037/a0026375

Arndt, J. (2015). The influence of forward and backward associative strength on false memories for encoding context. *Memory*, **23**(15), 1093–1111. doi: 10.1080/09658211.2014.959527

Arndt, J., & Beato, M. S. (2017). The role of language proficiency in producing false memories. *Journal of Memory and Language*, **95**, 146–158. doi: 10.1016/j.jml.2017.03.004

Arndt, J., & Gould, C. (2006). An examination of two-process theories of false recognition. *Memory*, **14**(7), 814–833. doi: 10.1080/09658210600680749.

Arndt, J., & Hirshman, E. (1998). True and false recognition in MINERVA2: Explanations from a global matching perspective. *Journal of Memory and Language*, **39**(3), 371–391. doi: 10.1006/jmla.1998.2581

Beato, M. S., & Arndt, J. (2014). False recognition production indexes in forward associative strength (FAS) lists with three critical words. *Psicothema*, **26**(4), 457–463. doi: 10.7334/psicothema2014.79

Beato, M. S., & Arndt, J. (2017). The role of backward associative strength in false recognition of DRM lists with multiple critical words. *Psicothema*, **29**(3), 358–363. doi: 10.7334/psicothema2016.248

Beato, M. S., & Arndt, J. (2021). The effect of language proficiency and associative strength on false memory. *Psychological Research*, **85**(8), 3134–3151. doi: 10.1007/s00426-020-01449-3

Beato, M. S., Suarez, M., & Cadavid, S. (2023). Disentangling the effects of backward/forward associative strength and theme identifiability in false memory. *Psicothema*. **35**(2), 178–188. doi: 10.7334/psicothema2022.288

Ben-Artzi, E., Faust, M., & Moeller, E. (2009). Hemispheric asymmetries in discourse processing: Evidence from false memories for lists and texts. *Neuropsychologia*, **47**(2), 430–438. doi: 10.1016/j.neuropsychologia.2008.09.021

Bialystok, E. (2017). The bilingual adaptation: How minds accommodate experience. *Psychological Bulletin*, **143**(3), 233–262. doi: 10.1037/bul0000099

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.

Boldini, A., Beato, M. S., & Cadavid, S. (2013). Modality-match effect in false recognition: An event-related potential study. *Neuroreport*, **24**(3), 108–113. doi: 10.1097/WNR.0b013e32835c93e3

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science*, **11**(5), 164–169. doi: 10.1111/1467-8721.00192

Brainerd, C. J., & Reyna, V. F. (2005). *The science of false memory*. Oxford University Press.

Cabeza, R., & Lennartson, E. R. (2005). False memory across languages: Implicit associative response vs fuzzy trace views. *Memory*, **13**(1), 1–5. doi: 10.1080/09658210344000161

Cadavid, S., & Beato, M. S. (2017). False recognition in DRM lists with low association: A normative study. *Psicologica*, **38**, 133–147. https://www.uv.es/psicologica/articulos1.17/6CADAVID.pdf

Cadavid, S., Beato, M. S., Suarez, M., & Albuquerque, P. B. (2021). Feelings of contrast at test reduce false memory in the Deese/Roediger-McDermott paradigm. *Frontiers in Psychology*, **12**, Article 686390. doi: 10.3389/FPSYG.2021.686390

Cann, D. R., McRae, K., & Katz, A. N. (2011). False recall in the Deese-Roediger-McDermott paradigm: The roles of gist and associative strength. *Quarterly Journal of Experimental Psychology*, **64**(8), 1515–1542. doi: 10.1080/17470218.2011.560272

Carneiro, P., & Fernandez, A. (2010). Age differences in the rejection of false memories: The effects of giving warning instructions and slowing the presentation rate. *Journal of Experimental Child Psychology*, **105**(1–2), 81–97. doi: 10.1016/j.jecp.2009.09.004

Chen, J. C. W., Li, W., Westerberg, C. E., & Tzeng, O. J.-L. (2008). Test-item sequence affects false memory formation: An event-related potential study. *Neuroscience Letters*, **431**(1), 51–56. doi: 10.1016/j.neulet.2007.11.020

Ciaramelli, E., Ghetti, S., Frattarelli, M., & Làdavas, E. (2006). When true memory availability promotes false memory: Evidence from confabulating patients. *Neuropsychologia*, **44**(10), 1866–1877. doi: 10.1016/j.neuropsychologia.2006.02.008

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, **58**(1), 17–22. doi: 10.1037/h0046671

Diekelmann, S., Born, J., & Wagner, U. (2010). Sleep enhances false memories depending on general memory performance. *Behavioural Brain Research*, **208**(2), 425–429. doi: 10.1016/j.bbr.2009.12.021

Diekelmann, S., Landolt, H.-P., Lahl, O., Born, J., & Wagner, U. (2008). Sleep loss produces false memories. *PLoS ONE*, **3**(10), Article e3512. doi: 10.1371/journal.pone.0003512

Dubuisson, J.-B., Fiori, N., & Nicolas, S. (2012). Repetition and spacing effects on true and false recognition in the DRM paradigm. *Scandinavian Journal of Psychology*, **53**(5), 382–389. doi: 10.1111/j.1467-9450.2012.00963.x

Fernandez, A., Diez, E., & Alonso, M. A. (2003). *Normas de Asociación libre en castellano, online database [Free Association norms in Spanish]*. Retrieved from: https://iblues-inico.usal.es/iblues/nalc_home.php

Gallo, D. A. (2006). *Associative illusions of memory: False memory research in DRM and related tasks*. Psychology Press.

Gallo, D. A. (2010). False memories and fantastic beliefs: 15 years of the DRM illusion. *Memory & Cognition*, **38**(7), 833–848. doi: 10.3758/MC.38.7.833

García-Bajos, E., & Migueles, M. (1997). Falsas memorias en el recuerdo y reconocimiento de palabras [False memories in recall and word recognition]. *Estudios de Psicología*, **18**(58), 3–14. doi: 10.1174/021093997320954818

Geng, H., Qi, Y., Li, Y., Fan, S., Wu, Y., & Zhu, Y. (2007). Neurophysiological correlates of memory illusion in both encoding and retrieval phases. *Brain Research*, **1136**(1), 154–168. doi: 10.1016/j.brainres.2006.12.027

Graves, D. F., & Altarriba, J. (2014). False memory in bilingual speakers. In R. R. Heredia & J. Altarriba (Eds.), *Foundations of Bilingual memory* (pp. 205–221). Springer. doi: 10.1007/978-1-4614-9218-4

Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, **9**(5), 220–225. doi: 10.1016/j.tics.2005.03.003

Howe, M. L., Gagnon, N., & Thouas, L. (2008). Development of false memories in bilingual children and adults. *Journal of Memory and Language*, **58**(3), 669–681. doi: 10.1016/j.jml.2007.09.001

Huff, M. J., Bodner, G. E., & Fawcett, J. M. (2014). Effects of distinctive encoding on correct and false memory: A meta-analytic review of costs and benefits and their origins in the DRM paradigm. *Psychonomic Bulletin & Review*, **22**(2), 349–365. doi: 10.3758/s13423-014-0648-8

Huff, M. J., Di Mauro, A., Coane, J. H., & O'Brien, L. M. (2021). Mapping the time course of semantic activation in mediated false memory: Immediate classification, naming, and recognition. *Quarterly Journal of Experimental Psychology*, **74**(3), 483–496. doi: 10.1177/1747021820965061

Iacullo, V. M., & Marucci, F. S. (2016). Normative data for Italian Deese/Roediger-McDermott lists. *Behavior Research Methods*, **48**(1), 381–389. doi: 10.3758/s13428-015-0582-3

Johansson, M., & Stenberg, G. (2002). Inducing and reducing false memories: A Swedish version of the Deese-Roediger-McDermott paradigm. *Scandinavian Journal of Psychology*, **43**(5), 369–383. doi: 10.1111/1467-9450.00305

Kawasaki, Y., & Yama, H. (2006). The difference between implicit and explicit associative processes at study in creating false memory in the DRM paradigm. *Memory*, **14**(1), 68–78. doi: 10.1080/09658210444000520

Kawasaki-Miyaji, Y., Inoue, T., & Yama, H. (2003). Cross-linguistic false recognition: How do Japanese-dominant bilinguals process two languages: Japanese and English? *Psychologia: An International Journal of Psychological Sciences*, **46**(4), 255–267. doi: 10.2117/psysoc.2003.255

Klimova, B. (2018). Learning a foreign language: A review on recent findings about its effect on the enhancement of cognitive functions among healthy older individuals. *Frontiers in Human Neuroscience*, **12**, Article 305. doi: 10.3389/fnhum.2018.00305

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**(2), 149–174. doi: 10.1006/jmla.1994.1008

Kroll, J. F., Van Hell, J. G., Tokowicz, N., & Green, D. W. (2010). The revised hierarchical model: A critical review and assessment. *Bilingualism: Language and Cognition*, **13**(3), 373–381. doi: 10.1017/S136672891000009X

Lee, Y.-S., Iao, L.-S., & Lin, C.-W. (2007). False memory and schizophrenia: evidence for gist memory impairment. *Psychological Medicine*, **37**(4), 559–567. doi: 10.1017/S0033291706009044

Lehtonen, M., Soveri, A., Laine, A., Järvenpää, J., de Bruin, A., & Antfolk, J. (2018). Is bilingualism associated with enhanced executive functioning in adults? A meta-analytic review. *Psychological Bulletin*, **144**(4), 394–425. doi: 10.1037/bul0000142

Mao, W. B., Yang, Z. L., & Wang, L. S. (2010). Modality effect in false recognition: Evidence from Chinese characters. *International Journal of Psychology*, **45**(1), 4–11. doi: 10.1080/00207590902757641

Marmolejo, G., Diliberto-Macaluso, K. A., & Altarriba, J. (2009). False memory in bilinguals: Does switching languages increase false memories? *The American Journal of Psychology*, **122**(1), 1–16. https://www.jstor.org/stable/27784371

McDermott, K. B., & Roediger, H. L. (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory and Language*, **39**(3), 508–520. doi: 10.1006/jmla.1998.2582

Nelson, D. L., McEvoy, C. L., Schreiber, T. A. (1998). *The University of South Florida word association, rhyme, and word fragment norms*. Retrieved from http://w3.usf.edu/FreeAssociation/

Plancher, G., Nicolas, S., & Piolino, P. (2008). Influence of suggestion in the DRM paradigm: What state of consciousness is associated with false memory? *Consciousness and Cognition*, **17**(4), 1114–1122. doi: 10.1016/j.concog.2008.08.006

Quinteros Baumgart, C., & Billick, S. B. (2018). Positive cognitive effects of bilingualism and multilingualism on cerebral function: A review. *Psychiatric Quarterly*, **89**(2), 273–283. doi: 10.1007/s11126-017-9532-9

Rocha, A. A. M., & Albuquerque, P. B. (2003). Ilusões de memória em alcoólicos [Memory illusions in alcoholics]. *Psicologia: Investigação e Prática*, **2**, 269–288.

Roediger, H. L., Balota, D. A., & Watson, J. M. (2001). Spreading activation and arousal of false memories. In H. L. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 95–115). American Psychological Association. doi: 10.1037/10394-006

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**(4), 803–814. doi: 10.1037/0278-7393.21.4.803

Rummer, R., Schweppe, J., & Martin, R. C. (2009). A modality congruency effect in verbal false memory. *European Journal of Cognitive Psychology*, **21**(4), 473–483. doi: 10.1080/09541440802188255

Sahlin, B. H., Harding, M. G., & Seamon, J. G. (2005). When do false memories cross language boundaries in English-Spanish bilinguals? *Memory & Cognition*, **33**(8), 1414–1421. http://www.ncbi.nlm.nih.gov/pubmed/16615389

Schacter, D. L., & Slotnick, S. D. (2004). The cognitive neuroscience of memory distortion. *Neuron*, **44**(1), 149–160. doi: 10.1016/j.neuron.2004.08.017

Stadler, M. A., Roediger, H. L., & McDermott, K. B. (1999). Norms for word lists that create false memories. *Memory & Cognition*, **27**(3), 494–500. doi: 10.3758/BF03211543

Stein, L. M., Feix da, L. F., & Rohenkohl, G. (2006). Avanços metodológicos no estudo das falsas memórias: Construção e normatização do procedimento de palavras associadas [Methodological advances in the study of false memories: Construction and standardization of the associated words procedure]. *Psicologia: Reflexão e Crítica*, **19**(2), 166–176. doi: 10.1590/S0102-79722006000200002

Stein, L. M., & Pergher, G. K. (2001). Criando falsas memórias em adultos por meio de palavras associadas [Creating false memories in adults by means of associated words]. *Psicologia: Reflexão e Crítica*, **14**(2), 353–366. doi: 10.1590/S0102-79722001000200010

Suarez, M., & Beato, M. S. (2021). The role of language proficiency in false memory: A mini review. *Frontiers in Psychology*, **12**, Article 659434. doi: 10.3389/FPSYG.2021.659434

Ulatowska, J., & Olszewska, J. (2013). Creating associative memory distortions – A Polish adaptation of the DRM paradigm. *Polish Psychological Bulletin*, **44**(4), 449–456. doi: 10.2478/ppb-2013-0048

Van Damme, I., & D'Ydewalle, G. (2009a). Memory loss versus memory distortion: The role of encoding and retrieval deficits in Korsakoff patients' false memories. *Memory*, **17**(4), 349–366. doi: 10.1080/09658210802680349

Van Damme, I., & D'Ydewalle, G. (2009b). Implicit false memory in the DRM paradigm: Effects of amnesia, encoding instructions, and encoding duration. *Neuropsychology*, **23**(5), 635–648. doi: 10.1037/a0016017

Wang, J., Otgaar, H., Howe, M. L., & Zhou, C. (2019). A self-reference false memory effect in the DRM paradigm: Evidence from Eastern and Western samples. *Memory & Cognition*, **47**(1), 76–86. doi: 10.3758/s13421-018-0851-3

Wang, J., Otgaar, H., Howe, M. L., Lippe, F., & Smeets, T. (2018). The nature and consequences of false memories for visual stimuli. *Journal of Memory and Language*, **101**, 124–135. doi: 10.1016/j.jml.2018.04.007

Watson, J. M., McDermott, K. B., & Balota, D. A. (2004). Attempting to avoid false memories in the Deese/Roediger-McDermott paradigm: assessing the combined influence of practice and warnings in young and old adults. *Memory & Cognition*, **32**(1), 135–141. doi: 10.3758/BF03195826

Yonelinas, A. P., Aly, M., Wang, W.-C., & Koen, J. D. (2010). Recollection and familiarity: Examining controversial assumptions and new directions. *Hippocampus*, **20**(11), 1178–1194. doi: 10.1002/hipo.20864

Zeelenberg, R., & Pecher, D. (2002). False memories and lexical decision: even twelve primes do not cause long-term semantic priming. *Acta Psychologica*, **109**(3), 269–284. doi: 10.1016/s0001-6918(01)00060-9

Appendix

The sixteen ten-word study lists with the original and translated critical lure and the language for each list

List	Language	Original critical lure	Translated critical lure	Associated words
List 1	English	MOUNTAIN	MONTAÑA	hill, valley, climb, summit, top, molehill, peak, plain, glacier, goat
List 2	English	MUSIC	MÚSICA	note, sound, piano, sing, radio, band, melody, horn, concert, instrument
List 3	English	FRUIT	FRUTA	apple, vegetable, orange, kiwi, citrus, ripe, pear, banana, berry, cherry
List 4	English	LION	LEÓN	tiger, circus, jungle, tamer, den, cub, Africa, mane, feline, roar
List 5	English	THIEF	LADRÓN	steal, robber, crook, burglar, money, cop, bad, rob, jail, gun
List 6	English	NEEDLE	AGUJA	thread, pin, eye, sewing, sharp, point, prick, thimble, haystack, thorn
List 7	English	SHIRT	CAMISA	blouse, sleeves, pants, tie, button, shorts, iron, polo, collar, vest
List 8	English	KING	REY	queen, England, crown, prince, George, dictator, palace, throne, chess, rule
List 9	Spanish	CINE	CINEMA	película, arte, televisión, oscuro, visión, actor, teatro, mudo, pantalla, espectáculo
List 10	Spanish	FLOR	FLOWER	primavera, olor, campo, rosa, margarita, amapola, pétalos, polen, capullo, aroma
List 11	Spanish	LÁMPARA	LAMP	luz, salón, techo, mesilla, Aladino, flexo, habitación, gas, magia, pared
List 12	Spanish	BOTELLA	BOTTLE	vino, agua, bebida, verde, alcohol, líquido, mensaje, vodka, tapón, vidrio
List 13	Spanish	CURA	PRIEST	iglesia, monja, sotana, misa, fraile, párroco, sacerdote, santa, religión, pueblo
List 14	Spanish	LIBRO	BOOK	leer, lectura, letras, hojas, estudiar, entretenimiento, aprender, sabiduría, página, estantería
List 15	Spanish	CAJA	BOX	dinero, fuerte, cartón, sorpresa, cajón, secreto, guardar, regalo, vacío, cuadrada
List 16	Spanish	SILLA	CHAIR	mesa, sillón, sentarse, descanso, asiento, cansancio, sofá, taburete, cocina, respaldo

Cite this article: Beato, MS., Albuquerque, PB., Cadavid, S., and Suarez, M. (2023). The effect of memory instructions on within- and between-language false memory. *Applied Psycholinguistics* **44**, 179–203. https://doi.org/10.1017/S0142716423000140