Bilingualism: Language and Cognition

cambridge.org/bil

Research Article

Cite this article: Elias M, Degani T (2022). Cross-language interactions during novel word learning: The contribution of form similarity and participant characteristics. *Bilingualism: Language and Cognition* **25**, 548–565. https://doi.org/10.1017/S1366728921000857

Received: 19 May 2020
Revised: 22 August 2021
Accepted: 22 August 2021
First published online: 20 January 2022

Keywords:
False cognates; vocabulary learning; individual differences; phonological short-term memory

Address for correspondence:
Tamar Degani, Department of Communication Sciences & Disorders, University of Haifa, Haifa, Israel
Email: tdegani@research.haifa.ac.il

Cross-language interactions during novel word learning: The contribution of form similarity and participant characteristics

Mariana Elias and Tamar Degani

University of Haifa, Israel

Abstract

The study examined whether false-cognates, overlapping in form but not meaning across languages, are easier to learn due to form overlap, or more difficult to learn due to meaning competition, compared to unambiguous control and cognate words. Fifty-four native Hebrew speakers learned 14 cognates, 14 false-cognates, and 28 control Arabic words in one session. Cognates were learned better than control items. There was no overall difference in learning false-cognates relative to controls, but individuals with higher phonological short-term memory, or with lower L1 verbal fluency, did exhibit a false-cognate learning-advantage. For these individuals, form overlap was more influential than meaning competition. Lexical decisions to Hebrew words following Arabic learning were slower for false-cognates than controls, indicative of backward influences. The findings reveal the influence of prior knowledge on learning and processing, and highlight the importance of jointly considering item-based and learner-based characteristics during the initial stages of vocabulary learning.

Introduction

Among the many challenges associated with learning a new language, learners need to acquire extensive new vocabulary to become proficient users of the language (e.g., Nation, 2013; Perfetti, 2007). Improving initial vocabulary acquisition and identifying the factors that may modulate it are therefore of great importance (Rice & Tokowicz, 2020; Tokowicz & Degani, 2015). In the current study, we focused on understanding how the overlap between the to-be-learned materials and participants' prior knowledge affect foreign-language (FL) learning. We examined how overlap in form and in meaning between the to-be-learned words and learners' first language (L1) affected the learning trajectories of different individuals. Whereas it is known that cognates that overlap in both form and meaning across languages (e.g., /ʔozen/ means 'ear' in both Arabic and Hebrew) are easier to learn compared to other types of words (Lotto & de Groot, 1998), it is less clear how form-overlap in the absence of meaning-overlap affects learning. Therefore, the aim of the current study was to examine how FALSE-COGNATES, which overlap in phonological and/or orthographic form across languages but differ in meaning (e.g., /sˤu:sˤ/ refers to a 'chick' in Arabic but to a 'horse' in Hebrew) are learned. Furthermore, we aimed to answer two additional questions: (1) how individual characteristics (phonological short-term memory and L1 verbal fluency) modulate the learning efficiency of this FL vocabulary, and (2) whether the newly learned words could have an influence on L1 processing.

The role of L1 in novel word learning

A point of departure for many models that aim to explain FL vocabulary learning is that learners rely on prior knowledge during learning. For example, a key feature of the Unified Competition Model (MacWhinney, 2005) is that, due to the interconnections between the L1 and the FL/second language (L2), whatever can transfer from the L1 will transfer to affect learning and processing of the FL/L2. The model proposes that in the lexical domain, the L2 is parasitic to the L1, and relies on the formal structure of the L1, at least at the beginning stages of learning. Similarly, according to Ecke's Parasitic Model (2015), when learning novel words, learners detect form similarities between the to-be-learned word and existing representations, such that the novel word is connected to a host representation in the previously known language(s). The Revised Hierarchical Model (RHM, Kroll & Stewart, 1994) postulates that with low L2 proficiency, L2 words rely on their L1 translations before developing independent links from form to meaning. Further, the Typological/Contrastive approach (Odlin, 1989) similarly suggests that FL learning capitalizes on knowledge from the L1, and postulates that a FL which is structurally similar to the L1 would be easier to learn than a FL which is not similar to it.

Accordingly, the similarity of the to-be-learned language to the L1 is expected to influence the learning outcome.

Of relevance, language similarity may be operationalized as language typological similarity (e.g., English and Italian vs. English and Arabic) resulting in manipulation of this variable at the whole language level. At the same time, language similarity may be manipulated at the item level, with particular items being more or less similar across languages. For instance, in a recent line of studies, the item-based overlap across languages, as reflected in the mapping of specific translations (i.e., translation ambiguity), was found to influence FL vocabulary learning (Degani & Tokowicz, 2010; Degani & Goldberg, 2019; Degani, Tseng & Tokowicz, 2014). Here, we extend this line of research to examine how semantic-form ambiguity (the mapping of a single form to different meanings, i.e., false-cognates) influences learning. In the current study, learners' L1 (Hebrew) and the FL they learned (Arabic) are linked at the whole language level, because both are Semitic languages that share partially similar phonological and morphological properties (Ibrahim, 2006; Saiegh-Haddad & Henkin-Roitfarb, 2014). However, we specifically focused on an item-based similarity manipulation, by including items that do or do not overlap in form across the two languages.

The role of semantic-form overlap in novel word learning

Cognate learning

One can think of the process of novel word learning as incorporating the learning of three dimensions: form, meaning and the mapping between these two (Tokowicz & Degani, 2015). Cognates overlap in all three dimensions across languages and thus may enjoy an advantage in FL vocabulary learning, because transfer from the L1 can occur smoothly without producing errors (MacWhinney, 2005). Indeed, Lotto and de Groot (1998) tested cognate and non-cognate word learning among adult Dutch speakers who learned Italian vocabulary. Participants showed better learning of cognate words, in both accuracy and latency measures, regardless of whether the words were learned in association with a Dutch translation word or a picture. Similarly, in a study conducted by Raboyeau, Marcotte, Adrover-Roig, and Ansaldo (2010), French speakers learned Spanish cognates better than non-cognates. This advantage was reflected in both an early learning phase (spanning over 5 days), and in a later consolidation period (of 2 weeks). A similar cognate facilitation effect was observed among learners of more distant languages (Persian and French, Ghazi-Saidi & Ansaldo, 2017).

False-cognate learning – form facilitation vs. meaning competition

In contrast, false-cognates (e.g., /sˤu:sˤ/) overlap in form but differ in meaning across languages. The form overlap may facilitate learning of the form component (Mulík, Carrasco-Ortiz & Amengual, 2019; Storkel, Maekawa & Aschenbrenner, 2013), but because the meanings do not overlap, the new mappings may impose difficulty in learning, due to the need to map a new unrelated meaning to the already existing and established word-form (Fang, Perfetti & Stafura, 2017; Maciejewski, Rodd, Mon-Williams & Klepousniotou, 2020; Rodd, Berriman, Landau, Lee, Ho, Gaskell & Davis, 2012).

Previous studies that examined learning of new meanings to a known form within a language may shed light on how false-cognates are learned. Specifically, learning a new meaning to a known word-form creates an ambiguous word, in that a single form is now mapped onto two different meanings. Storkel et al. (2013) examined such learning amongst children and observed an advantage for learning new meanings to known forms compared to learning completely new words, attributed to better access to the frequent known word-forms. Thus, word-form familiarity may facilitate learning.

At the same time, other research suggests that learning an unrelated meaning to an already known form may hinder learning. For instance, Rodd et al. (2012) taught adult native English monolinguals fictional new meanings for unambiguous English words and showed better learning for meanings that were semantically related to the original meaning (e.g., Hive as a busy household), compared to those that were semantically unrelated meanings (e.g., Hive as a mythical monster). This latter finding indicates that learning a new unrelated meaning may impose a challenge (see also Bracken, Degani, Eddington & Tokowicz, 2017; Floyd & Goldberg, 2020; Maciejewski et al., 2020).

Together, these two lines of research may reveal the operation of non-mutually exclusive processes: namely, form facilitation (Storkel et al., 2013) and meaning competition (Rodd et al., 2012), which may affect learning performance. In support of this possibility, recent work provides evidence for both form facilitation and meaning competition, taking place at different time points in learning. Specifically, Fang et al. (2017) compared learning novel words (new forms and meanings) to learning new meanings to already known forms by examining learning efficiency, indexed as the ability to generate meanings relative to the time spent studying the word. Interestingly, a biphasic effect emerged in that in the early stages of learning participants were more efficient in learning new meanings to already known forms, but over time the pattern reversed with a greater learning efficiency for the novel words.

Therefore, the extent to which overlap in form can facilitate learning when there is no meaning-overlap remains unclear. Moreover, the majority of past research tested learning of new meanings to already existent forms WITHIN a given language (but see Ghazi-Saidi & Ansaldo, 2017; Mulík et al., 2019; Otwinowska, Foryś-Nogala, Kobosko & Szewczyk, 2020), whereas our study aims to better understand the role of phonological-overlap, with and without a semantic-overlap, ACROSS the language boundary in two different languages. One relevant study conducted recently by Mulík et al. (2019), examined how Spanish–English bilinguals with low and high L2 proficiency learned three types of Slovak words: overlapping in phonology but not in meaning with the L1 (Spanish false-cognates), overlapping in phonology but not in meaning with the L2 (English false-cognates) and control words, not overlapping with either of the languages. The results indicated better learning of false-cognates of both types in comparison to control words in the high proficiency group, and better learning of L1 false-cognates by the low proficiency group.

Contrasting findings, suggesting worse learning of false-cognates relative to control, were also observed. Specifically, in a recent study, Otwinowska and Szewczyk (2019) examined the learnability of different L2 words by asking adult Polish learners of English to provide the translation of a list of English words and rate their confidence in the translation. Interestingly, false-cognates were found to be less learnable than control and cognate words, although learners presumably had comparable opportunities to learn these words in the input.

Other research observed no difference between false-cognates and control words (Ghazi-Saidi & Ansaldo, 2017; Otwinowska et al., 2020). In particular, Ghazi-Saidi and Ansaldo (2017) taught adult Persian speakers cognates, false-cognates and control French words as novel words. Results showed faster naming of cognates in comparison to false-cognates and control words, with no significant difference in naming latencies between false-cognates and control words. At the same time, false-cognates and cognates differed from control words in the recruitment of certain brain areas (basal gangilia and parahippocampal gyrus) thought to reflect the engagement of implicit memory systems, suggesting that the similar learning accuracy for false-cognates and controls may be based on different mechanisms (see also Grant, Fang & Li, 2015).

Given the contradicting findings of the previous studies and the lack of certainty regarding false-cognate learning, we aimed to provide additional insight into the way in which false-cognates are learned in comparison to control words. Therefore, participants in the current study learned Arabic false-cognates with meanings that are semantically unrelated to the original Hebrew words. Based on the results of Mulík et al. (2019) and Storkel et al. (2013), facilitation may be expected for false-cognates over control words due to the overlap in form. In contrast, according to the findings of Otwinowska and Szewczyk (2019) and Rodd et al. (2012), we expected a challenge in learning these words in comparison to unambiguous control words due to the difficulty in mapping unrelated meanings. Further, to better understand the conflicting pattern that emerged from previous research, we suggest that participants' individual characteristics should be considered. We propose that learners' cognitive and linguistic profile may modulate the extent to which they benefit from form facilitation or suffer from meaning competition during learning. Therefore, in the current study, interactions between item-based and learner-based characteristics were examined.

Individual differences

Phonological short-term memory

Our study further aimed to test how individual characteristics modulate novel word learning. Specifically, better phonological short-term memory has been linked to better vocabulary learning of foreign and artificial languages both in adults (Cheung, 1996; Kaushanskaya, 2012; Martin & Ellis, 2012) and in children (Gathercole, Hitch & Martin, 1997; Michas & Henry, 1994). Interestingly, Degani and Goldberg (2019) found that better phonological short-term memory was associated with better overall FL vocabulary learning, yet it was also associated with a greater disadvantage in learning translation-ambiguous words. Presumably, the stronger activation of phonological forms for individuals with higher phonological short-term memory resulted in increased competition in the ambiguous condition in which two different words in the FL (Arabic) corresponded to one form in L1 (Hebrew). In the current study, we expected to observe an overall advantage for individuals with higher phonological short-term memory as measured by a non-word repetition task (Shatil & Share, 2003). We further tested whether it modulated the difference in learning false-cognates vs. control words. We expected individuals with higher phonological short-term memory abilities to capitalize on the phonological overlap of false-cognates, such that for these individuals, false-cognates would be easier to learn compared to control words.

L1 abilities

FL vocabulary learning has also been positively linked with L1 proficiency (Geva, 2014). For instance, Sparks, Patton, Ganschow, Humbach, and Javorsky (2006) showed that literacy measures in the L1 predicted better FL proficiency 10 years later. Mulík et al. (2019) demonstrated that participants with higher L2 proficiency were better learners of false-cognates in an L3 in comparison to participants with lower L2 proficiency. Notably, Degani and Goldberg (2019) found that increased L1 verbal fluency, rather than subjective ratings of proficiency in the language from which participants learned the FL vocabulary, was associated with better learning. Although L1 verbal fluency could be thought of as a subcomponent of the complex construct of language proficiency (Hulstijn, 2011), in the Degani and Goldberg (2019) sample there was no correlation between the subjective ratings and the verbal fluency score. Verbal fluency may reflect cognitive control in addition to lexical access (Carpenter, Rao, Peñaloza & Kiran, 2020; Friesen, Luo, Luk & Bialystok, 2015), but critically it may index the availability of L1 representations at the time of learning (Linck, Kroll & Sunderman, 2009). Therefore, and because self-ratings of proficiency may be of limited range and predictive power in the L1 (Tomoschuk, Ferreira & Gollan, 2019), in the current study we focused on participants' L1 (Hebrew) verbal fluency score, as measured by a semantic fluency task (Kavé, 2005), as an index of the availability of learners' L1 representations. We further examined individual differences in current patterns of L1 (Hebrew) use, as these may similarly affect the availability of L1 representations during learning.

Notably, the link between knowledge of prior languages and novel vocabulary learning may stem from two reasons (see also Hirosh & Degani, 2018). The first is the reliance of individuals on direct transfer from their previous languages, in that individuals with higher abilities in the known languages can more easily access phonological and semantic representations in their lexicon (Gathercole, 2006). The second is that observed L1 verbal fluency in fact reflects individuals' general language-learning capacity (a common proficiency/aptitude construct, e.g., Cummins, 1979, 1991), which may be linked to more general cognitive abilities (Geva, 2014). Under both accounts, we expected learners with better L1 verbal fluency to be overall better at FL vocabulary learning.

Further, the degree to which L1 verbal fluency modulates how different word types are learned may take several forms. To the extent that the advantage of individuals with higher L1 verbal fluency stems from their specific reliance on L1 phonological and semantic representations in their lexicon (Gathercole, 2006), we expected individuals with higher L1 verbal fluency to show a larger advantage in learning novel words with semantic-form overlap (cognates). Reliance on L1 meaning may lead individuals with higher L1 abilities to show a false-cognate disadvantage relative to controls, given the misalignment in form to meaning mapping for false-cognates. However, to the extent that the effect of L1 verbal fluency comes from a common proficiency construct, the difference between individuals with lower and higher abilities may become larger under more demanding learning conditions: namely, false-cognates and controls but not cognates.

Bidirectional cross-language effects

The final goal of the present study was to investigate whether cross-language interactions that are often observed for proficient

https://doi.org/10.1017/S1366728921000857 Published online by Cambridge University Press

bilinguals can be found at the initial stages of learning. In particular, proficient bilinguals exhibit cross-language interactions when presented with words with semantic-form overlap across their two languages (Degani, Prior & Hajajra, 2018; Prior, Degani, Awawdy, Yassin & Korem, 2017). For example, Degani et al. (2018) observed that proficient Arabic–Hebrew bilinguals activated the Arabic meaning of Hebrew/Arabic cognates and false-cognates during a visual semantic-relatedness judgment task in Hebrew (see also Prior et al., 2017 for auditory presentation). This cross-language influence was present even though bilinguals performed a task in which the other language was not relevant, suggesting that activation of a word-form in one language leads to form and meaning activation in both languages (i.e., NON-SELECTIVE ACCESS, Dijkstra & van Heuven, 2002; Kroll, Bobb & Wodniecka, 2006). Thus, when individuals acquire new meanings for an already existent unambiguous word in their other language, as in the case of false-cognates in the current study, competition might occur between the original dominant meaning in the L1 and the newly learned meaning (see also Dumay & Gaskell, 2007, 2012; Rodd et al., 2012).

This competition may take the form of FORWARD influences across-languages, in that the L1 affects FL vocabulary learning. Notably, the current study further tested for BACKWARD cross-language influences, testing whether the newly learned meanings affect processing of the already existent forms in the L1. Such bidirectional cross-language influences have been demonstrated among more proficient bilinguals (Degani, Prior & Tokowicz, 2011; Malt, Jobe, Li, Pavlenko & Ameel, 2016). Further, in the study by Rodd et al. (2012) described above, learners responded more quickly to the words to which they learned a semantically related new meaning in comparison to unrelated new meanings in a semantic decision task following learning (see also Maciejewski et al., 2020).

These findings suggest the possibility that a newly learned meaning would affect processing of the established word form in the L1. To test this, following learning, participants performed a visual lexical decision task in their L1, containing Hebrew words that correspond to the newly learned Arabic cognates and false-cognates. To the extent that the newly learned Arabic cognates and false-cognates were integrated into a unified bilingual lexicon (Dijkstra & van Heuven, 2002), we expected that following learning, the Arabic meaning would be activated when learners were presented with the Hebrew existing form. Thus, we expected learners to respond to Hebrew false-cognates as ambiguous Hebrew words entailing more than one meaning. Conversely, we expected the Hebrew cognates to enjoy an advantage in processing due to the convergence of meanings from both languages.

The current study

To summarize, in the current study, native Hebrew speakers with no previous knowledge of Arabic learned three types of Arabic words in the spoken modality (cognates, false-cognates, and controls). All participants completed a set of training and testing cycles, as well as objective (a non-word repetition task to test phonological short-term memory, and an L1 verbal fluency task) and subjective (language history questionnaire) measures of linguistic and cognitive abilities. The study aimed to answer three main questions. First, we tested how semantic-form overlap affected learning by incorporating translation recognition tests intermitted with learning cycles. Thus, in the course of learning and testing we examined which of the three word types was learned more easily, and specifically whether cognates were learned better than unambiguous controls and whether false-cognates were harder or easier to learn compared to unambiguous control words. Second, we examined how individual differences modulated learning of FL vocabulary, with a focus on phonological short-term memory and L1 verbal fluency. The availability of L1 representations was also examined via ratings of current L1 use. Finally, we tested for backward cross language influences. Thus, we asked whether the newly learned cognates and false-cognates affected the processing of these forms in the L1 by employing an L1 visual lexical decision task post learning.

Method

Participants

Fifty-six native Hebrew speakers, with no or limited knowledge of Arabic, participated in the experiment. Thirty-four of them indicated learning Arabic in high school, but reported minimal proficiency and use of Arabic[1]. All participants had acquired English as an L2 or L3 and use it frequently in their academic or professional life. They were recruited from a large university in Israel, and indicated no learning or attention disabilities. Participants' language profile was assessed via objective (verbal fluency task in Hebrew) and subjective (language history questionnaire) language measures. Participants signed an informed consent as an approval of their participation, and were compensated with class credit or payment.

Data from two participants were excluded, one because Hebrew was not their native language, and the other due to incompletion of the protocol due to extremely long completion time of the learning phase. Analyses of word learning are therefore based on a final set of 54 participants (40 females, 14 males, average age 24.85). See Table 1 for more information. Notably, as discussed below, each cycle of learning was analyzed with the maximum number of participants who completed that cycle (54 in cycle 1; 37 in cycle 2; 22 in cycle 3, and 12 in cycle 4). Further, analysis of backward influences was based on a subset of 43 participants who reached a learning criterion of 80% in the learning phase.

Stimuli

Fifty-six Arabic words were taught to each participant, of which 14 words were cognates (e.g., /ʔozen/ which means 'ear' in both languages), 14 were false-cognates (e.g., /s^{ʕ}u:s^ʕ/ which means 'chick' in Arabic but 'horse' in Hebrew) and 28 were unambiguous control words. See Table 2 for mean characteristics and Supplementary Materials (Supplementary Materials, S1) for the full stimuli set. All Arabic stimuli were recorded in standard Arabic by a female native Arabic speaker.

To determine the phonological similarity of the Hebrew and Arabic forms, norming procedures were conducted such that native Hebrew speakers (with no previous knowledge of Arabic) rated the similarity of the aural form of the Arabic word along with the phonological form of the visually presented Hebrew word. Each item was rated by at least 10 participants on a scale

[1]To verify that this minimal Arabic exposure did not affect the results we conducted the analyses on the subset of participants who reported learning Arabic in high-school ($n = 34$). The pattern of results was identical, with better learning of cognates relative to control, and no difference between false-cognates and control in each cycle of learning, and across cycles.

https://doi.org/10.1017/S1366728921000857 Published online by Cambridge University Press

Table 1. Participants background information.

Measure	Average (SD)
Number of Participants	54
Age (in years)	24.85 (3.82)
Percent of Current Hebrew Exposure	82.22 (13.72)
Percent of Current Hebrew Reading	89.22 (12.37)
Percent of Current Hebrew Talking	86.20 (19.62)
Avg. of L1 (Hebrew) Proficiency[a]	9.44 (0.68)
Avg. of L1 (Hebrew) Use[b]	7.71 (1.23)
Avg. of L2* Proficiency[a]	6.85 (1.40)
Avg. of L2* Use[b]	5.63 (2.13)
Avg. of Arabic Proficiency[a]	0.66 (0.82)
Avg. of Arabic Use[b]	0.18 (0.39)
Number of Languages	2.57 (0.81)
Phonological Short-term Memory (range 0-14)	5.77 (1.36)
L1 (Hebrew) Verbal Fluency (average per minute)	16.98 (3.83)

Note: L2* is the second most proficient language reported by the participants, following Hebrew. It varied across participants, but was never Arabic, and was English for 47 of the participants. [a]Language proficiency reflects the average proficiency ratings in reading, writing, talking, and speech comprehension rated on a scale of 0 (lowest) to 10 (highest). [b]Language use reflects the average use in reading, writing, conversation, internet, listening, and TV watching rated on a scale from 0 (lowest) to 10 (highest).

Table 2. Mean item characteristics as a function of word type.

	Word Type		
Measure	Cognates	False-Cognates	Control
Hebrew length in letters	3.71 (0.91)	4.28 (0.99)	3.82 (0.98)
Hebrew length in syllables	2.00 (0.39)	2.07 (0.47)	2.11 (0.68)
Hebrew frequency	32.64 (31.44)	45.26 (42.96)	39.96 (56.64)
Arabic length in syllables	1.78 (0.57)	1.5 (0.51)	1.92 (0.85)

Note: There were no significant differences among the word types based on a one-way Anova in any of the measures (all p_s > 0.20). Hebrew frequency is based on heTenTen 2014 corpus via SketchEngine (Kilgarriff et al., 2014; Kilgarriff et al., 2010). Standard deviations appear in parenthesis.

of 1-5 (1 = minimum form similarity, 5 = maximum form similarity). All selected cognates and false-cognates were rated higher than 3 in phonological similarity between their Hebrew and Arabic forms, such that cognates ($M = 4.23$, $SD = 0.61$) did not differ significantly from false-cognates ($M = 4.36$, $SD = 0.53$), $t < 1$. Note, that although the cognates and false-cognates used in the current study substantially overlap in phonological form across Hebrew and Arabic, their phonological realization is not identical. Further, phonological similarity ratings of the Arabic false-cognates with the form of the appropriate Hebrew translations (e.g., /sˤuːsˤ/ in Arabic vs. /ʔefroaħ/ in Hebrew, both meaning ‘chick’) were lower than 3 ($M = 1.18$, $SD = .18$).

Selected words were matched across word types, such that cognates, false-cognates and control words did not differ significantly in Hebrew length in letters ($F(2, 55) = 1.45$, $MSE = .94$, $p = 0.24$), Hebrew length in syllables ($F < 1$), Hebrew frequency ($F < 1$, based on heTenTen 2014 corpus via SketchEngine, see Kilgarriff, Baisa, Bušta, Jakubíček, Kovář, Michelfeit, Rychlý & Suchomel, 2014; Kilgarriff, Reddy, Pomikálek & Avinesh, 2010), and Arabic length in syllables ($F(2, 55) = 1.64$, $MSE = .52$, $p = 0.20$). Further, items were matched on semantic density ($F < 1$) computed for the corresponding English translation (Elexicon, Balota, Yap, Cortese, Hutchison, Kessler, Loftis, Neely, Nelson, Simpson & Treiman, 2007) because there was no available information on semantic density in Hebrew or Arabic. Importantly, Hebrew translations differed significantly in phonological similarity ($F(2,55) = 311.14$, $MSE = .15$, $p < .001$). Planned comparisons with Bonferroni corrections revealed that items were significantly ($ps < .001$) more phonologically similar to their Arabic translations in the cognate word type ($M = 4.23$, $SD = .61$) than in the false-cognate word type ($M = 1.18$, $SD = .18$) and control word type ($M = 1.28$, $SD = .31$), which did not differ from each other ($p = 1.0$).

Procedure

Each participant completed the protocol in one session in which they completed up to four computerized learning cycles and translation recognition tests, all using E-prime software (Psychology Software Tools, Pittsburgh, PA). At the beginning of each learning or testing task, participants completed eight practice trials to get familiarized with the task, in the presence of the experimenter. Following the word learning and testing tasks, participants completed additional linguistic and cognitive tasks and a language history questionnaire.

Learning and testing procedure

Familiarization learning cycle

In the first cycle, participants were aurally presented with the to-be-learned Arabic words and visually presented with their Hebrew translation to provide their meaning, and were requested to repeat out loud each Arabic word they hear, with no feedback. Specifically, on each trial a 1000 ms fixation cross was presented followed by a 500 ms blank screen. A Hebrew word was then visually presented in its written form for 1000 ms, followed by an aural presentation of the to-be-learned Arabic translation via headphones. Subsequently, a question mark appeared on the screen until the participant’s vocal response. A blank screen of 1500 ms was presented before the next trial.

Second and subsequent learning cycles – retrieval attempt

In these cycles, participants were presented with the Hebrew words and were requested to attempt to produce their Arabic translations. Retrieval-based learning (Karpicke, 2012) was incorporated because practice testing has been shown to be a more efficient learning strategy in comparison to rehearsal or imitation (Kang, Gollan & Pashler, 2013; Rice & Tokowicz, 2020; Tokowicz & Degani, 2015). On each trial, following the 1000 ms fixation cross and the 500 ms blank screen, participants were visually presented with a Hebrew word for 1000 ms, followed by a question mark that appeared on the screen until the participant’s vocal response triggered the voice key. Following this retrieval attempt, a 1000 ms blank screen appeared followed by aural presentation of the correct Arabic translation, regardless of the participant’s answer (Kang et al., 2013), to keep the number of presentations of the Arabic word equal for all participants. An additional 1500 ms blank screen then appeared before the next trial.

https://doi.org/10.1017/S1366728921000857 Published online by Cambridge University Press

Recognition test

Participants then completed a recognition test, in which a 1000 ms fixation cross appeared on the screen followed by a 500 ms blank screen. Participants then heard an Arabic word via headphones and four Hebrew alternative translations appeared on the screen. Participants selected their response by button press (1 to 4), received feedback (correct or incorrect), and were visually presented with the correct Hebrew response, regardless of their accuracy on the trial, to strengthen learning and keep the number of presentations equal (Butler, Karpicke & Roediger, 2008; Karpicke & Roediger, 2008). The four Hebrew alternative translations were selected randomly on each trial from the pool of Hebrew translations used during the learning cycles.

To ensure that the novel FL words were sufficiently learned before backward influences of FL on L1 were examined, a learning criterion was used (Bartolotti & Marian, 2012). Success rates on the recognition test were calculated by the computer program. If the participant reached 80% success rate on all three types of words (cognate, false-cognate and control), the learning protocol was over. When performance was below 80% success rate in one (or all) of the word types, learning resumed using the retrieval attempt protocol, followed by a recognition test. Learning and testing were repeated until an 80% criterion was met in all 3 word types for a maximum of 4 cycles.

Individual differences measures

Phonological short-term memory task

A Hebrew non-word repetition task (Shatil & Share, 2003) was used in order to test participants' phonological short-term memory span. Although this task was developed for school age children (up to 6 years of age), it was successfully used with adult learners in previous studies (e.g., Degani & Goldberg, 2019). Participants were asked to repeat out loud sets of growing lengths of Hebrew non-words (2 to 8 characters, 2 words from each length) in the same order presented to them by the experimenter. The test ended when participants failed to accurately repeat both sets of the same length.

L1 (Hebrew) verbal fluency task

Participants were asked to name as many words as they can in Hebrew within one minute for each of two different semantic categories (Kavé, 2005; see also Gollan, Montoya & Werner, 2002). Two categories (fruits and vegetables, and vehicles) were each presented on a computer screen, followed by an hour glass marking the time limit for the task (60 seconds per category).

Language History Questionnaire (LHQ)

Participants completed an LHQ (adapted from the LEAP-Q questionnaire, Marian, Blumenfeld & Kaushanskaya, 2007) providing details regarding proficiency and use patterns in the languages they know.

FL on L1 influence

In order to test whether the newly learned words affected participants' L1 processing, participants completed a HEBREW VISUAL LEXICAL DECISION TASK. The Hebrew words included the 14 cognate words that were learned during the learning protocol, as well as 14 Hebrew false-cognates for which an Arabic false-cognate had been learned. For instance, in the learning protocol participants learned that /sˤu:sˤ'/ in Arabic means 'chick', whereas in the lexical decision task, the Hebrew word /sus/, meaning 'horse', was presented. In addition, control items learned during the learning phase were included in the lexical decision task as a comparison, but due to matching considerations (in that now the task included the Hebrew false-cognate items), only 24 of the original 28 control items were included. Across the three word types (cognates, false-cognates and control), items were matched on Hebrew length in letters and syllables, in Hebrew frequency and in Hebrew bigram and trigram frequency (all *ps*>.31, see Supplementary Materials (Supplementary Materials, S2) for summary characteristics). Thirty-two additional filler items (18 unambiguous and 14 ambiguous Hebrew words such as /mapa/ which means both a tablecloth and a map) were included in order to conceal the purpose of this task. In total, the task included 84 words and 84 orthographically legal non-words. The selected non-words were matched to the real words on length in number of letters and syllables, and in bigram and trigram Hebrew frequency (from a corpus of 12 million words from a collection of articles from the Hebrew newspaper *Haaretz,* see also Peleg, Degani, Raziq & Taha, 2020; see Supplementary Materials, S2 for the full analyses data), to avoid fast and superficial lexical decisions based on orthographic information only (Keuleers & Brysbaert, 2010). On each trial, a 500 ms fixation cross appeared, followed by the letter string. Participants were requested to decide whether the letter string is a real Hebrew word by clicking "yes" or "no" on the response box as quickly and as accurately as possible. The letter string remained on the screen until participants' response, or up to 4 seconds, at which point a fixation cross appeared to signal the next trial.

To verify that the results of the learners reflect the effect of Arabic learning, their performance on this Hebrew lexical decision task was compared to that of a second group of 30 native Hebrew speakers who did not participate in the learning paradigm (see Supplementary Materials (Supplementary Materials, S6) for detailed background information). The two groups were comparable except that the non-learners performed only the lexical decision task and completed the LHQ.

Results

In what follows, we first report on the analyses examining how word type (cognate, false-cognate, control words) affects learning, including (a) the number of cycles participants needed in order to reach an 80% learning criterion in each word type, followed by detailed analyses of the (b) error rates and (c) RTs within each cycle of learning (and across cycles). Next, we report the analyses of how individual differences in linguistic and cognitive abilities modulate learning. Finally, we report how lexical decisions to L1 (Hebrew) words are affected by prior learning of related vocabulary in the FL (Arabic).

Number of cycles to criterion

For each participant, we computed the cycle at which they first reached the 80% criterion for each word type. On average, participants reached the overall learning criterion with $M = 1.4$ cycles ($SD = .74$)[2]. Interestingly, a repeated measures Anova with

[2]Ten participants did not reach this criterion in one or more of the word types within the four cycles provided. These 10 participants were included in the relevant learning analyses described below, but not in the subsequent backward influences analyses which assume successful learning (i.e., that the learning criterion had been met).

Table 3. Error rate and RT for each word type per cycle.

Cycle	N	Word type: Cognate	False Cognates	Control
		Error Rate		
1	54	0.03 (0.01)	0.20 (0.02)	0.23 (0.01)
2	37	0.01 (0.01)	0.17 (0.02)	0.18 (0.01)
3	22	0.04 (0.01)	0.14 (0.02)	0.18 (0.02)
4	12	0.06 (0.02)	0.23 (0.04)	0.26 (0.03)
		RT (in ms)		
1	54	1305 (29)	2207 (59)	2322 (42)
2	37	1191 (30)	2030 (68)	2151 (45)
3	22	1050 (37)	1886 (85)	1982 (61)
4	12	1238 (55)	2057 (116)	1966 (78)

Note: N indicates the number of participants per cycle. Standard errors appear in parentheses.

Bonferroni corrections for multiple comparisons revealed that this varied by Word Type ($F(1.4, 74) = 7.79$, $MSE = .94$, $p = 0.003$, $\eta_p^2 = .13$), such that the criterion for Cognates was reached more quickly ($M = 1.04$, $SD = .19$ cycles) than that of the Control words ($M = 1.39$, $SD = .90$ cycles), and the False-Cognates ($M = 1.65$, $SD = .97$ cycles), which did not differ from each other.

Performance per cycle

Table 3 provides mean Error Rate and Response Times (RTs) for correct responses by Word Type (Cognate, False-Cognate, and Control) per Cycle.

Data analysis and model structure

Results were analyzed using linear mixed-effects models as implemented in the lme4 (Baayen, Davidson & Bates, 2008) and lmerTest packages (Kuznetsova, Brockhoff & Christensen, 2017) in R (Version 4.0.3, R Core Team, 2020). RTs were analyzed for correct responses only. A preliminary inspection of the RT data revealed trials with absurdly long RTs (above 40 seconds) in the translation recognition test, that most likely do not reflect relevant task processes. Thus, we excluded trials with RTs longer that 6.5 seconds, resulting in the exclusion of less than 2.5% of the data. Further, examination of the RT distribution revealed substantial deviation from normality. To improve skew and kurtosis, we explored log and inverse transformations. Based on the QQ plots, the skew (raw data:1.2; log transformation: 0.2; inverse transformation: -1.12), kurtosis (raw data:1.75; log transformation: -0.58; inverse transformation: 1.46), and Anderson-Darling Normality test (raw data: $A = 232.94$, $p < .001$; log transformation: $A = 18.09$, $p < .001$; inverse transformation: $A = 98.63$, $p < .001$), the log transformation provided a better remedy, and was thus adopted. Error rate data were analyzed following the binomial distribution in logistic mixed-effect models. Each cycle was analyzed separately with the maximum number of participants who completed that cycle (54 in cycle 1; 37 in cycle 2; 22 in cycle 3, and 12 in cycle 4), and additional analyses compared performance across cycles for the 22 participants who completed the first three cycles.

Within each cycle and each measure (log RTs; Error rate) the models included the fixed effect of Word Type, dummy (treatment) coded with Control items set as the reference, to allow comparisons of Cognates vs. Controls and False-Cognates vs. Controls[3]. A maximal model including by-participant and by-item intercepts and a by-participant slope for Word Type (Bell, Fairbrother & Jones, 2019; Brauer & Curtin, 2018) was submitted to a *buildmer* function in the *buildmer* package (v. 1.3, Voeten, 2019, as used in e.g., Johns & Steuck, 2021) in R (version 4.0.3, R Core Team, 2020), which uses the (g)lmer function from the lme4 package (v 1.1.-21, Bates, Maechler, Bolker & Walker, 2015). Starting from the maximal model, and using backward-fitting model selection procedure, the *buildmer* function systematically simplifies the random slopes until convergence in addition to using likelihood ratio tests (LRTs) to examine the contribution of random slopes to the fit of the model (one of the common methods to test model fit, Matuschek, Kliegl, Vasishth, Baayen & Bates, 2017, p. 308). This systematic procedure is not based on decisions made by the researcher (e.g., avoids the need to determine the cutoff point for component weights selection in a PCA on the random structure, Bates, Kliegl, Vasishth & Baayen, 2015), and is fully replicable from the data. The *buildmer* function further tests the contribution of each fixed effect to the model fit via a chi-squared test on the residual sum of squares of each model. P-values for all fixed effects were determined based on Satterthwaite degrees of freedom using the *lmerTest* package (v. 3.1-0, Kuznetsova et al., 2017), or the Wald degrees of freedom for binomial distribution. When necessary, to probe interactions and examine pairwise comparisons, the selected model was refitted using (g)lmer, and followed by the *testInteractions* function from the *phia* package (v. 0.2-1, Martinez, 2015) or with the *contrast* function from the *emmeans* package (v. 1. 5. 2-1, Lenth, 2020) with Bonferroni adjustments of multiple comparisons. In what follows, significance of main effects was derived from the *anova* function for each model (see Supplementary Materials, S3). Detailed models are presented in the Tables following the *summary* function. Intraclass correlations were computed using the *icc* function in the *performance* package (Lüdecke, Makowski, Waggoner & Patil, 2020).

Learning

Word type effect

In all cycles, in both error rate ($F_{\mathrm{Cycle1}}(2) = 16.86$, $p < .001$; $F_{\mathrm{Cycle2}}(2) = 14.30$, $p < .001$; $F_{\mathrm{Cycle3}}(2) = 8.74$, $p < .001$; $F_{\mathrm{Cycle4}}(2) = 7.25$, $p < .001$) and RT ($F_{\mathrm{Cycle1}}(2{,}58.11) = 23.66$, $p < .001$; $F_{\mathrm{Cycle2}}(2{,}59.53) = 27.82$, $p < .001$; $F_{\mathrm{Cycle3}}(2{,}54.67) = 27.49$, $p < .001$; $F_{\mathrm{Cycle4}}(2{,}52.47) = 16.69$, $p < .001$), there was a significant Word Type effect such that Cognates were recognized with less errors and faster than Control words, but False-Cognates did not differ

[3]The difference between Cognate and False-Cognate items was not of theoretical interest in the current study. Nonetheless, we examined this difference using the *contrast* function from the *emmeans* package (v. 1.5.3, Lenth, 2020). Translation recognition of Cognates was associated with less errors (Cycle 1: *Odds Ratio* = 0.08, $SE = 0.04$, $z = -4.67$, $p < .001$; Cycle 2: *Odds Ratio* = 0.05, $SE = 0.03$, $z = -4.73$, $p < .001$; Cycle 3: *Odds Ratio* = 0.23, $SE = 0.12$, $z = -2.93$, $p = 0.01$; Cycle 4: *Odds Ratio* = 0.18, $SE = 0.11$, $z = -2.88$, $p = 0.01$; Across Cycles: *Odds Ratio* = 0.10, $SE = 0.05$, $z = -5.11$, $p < .001$) and with faster responses (Cycle 1: $b = -0.50$ $SE = 0.10$, $df = 61.0$, $t = -4.94$, $p < .001$; Cycle 2: $b = -0.47$ $SE = 0.10$, $df = 64.3$, $t = -5.01$, $p < .001$; Cycle 3: $b = -0.52$ $SE = 0.09$, $df = 56.1$, $t = -5.66$, $p < .001$; Cycle 4: $b = -0.52$ $SE = 0.10$, $df = 56.0$, $t = -5.02$, $p < .001$; Across Cycles: $b = -0.49$ $SE = 0.08$, $t = -6.03$, $p < .001$) compared to False-Cognate items in all cycles of learning, and across cycles.

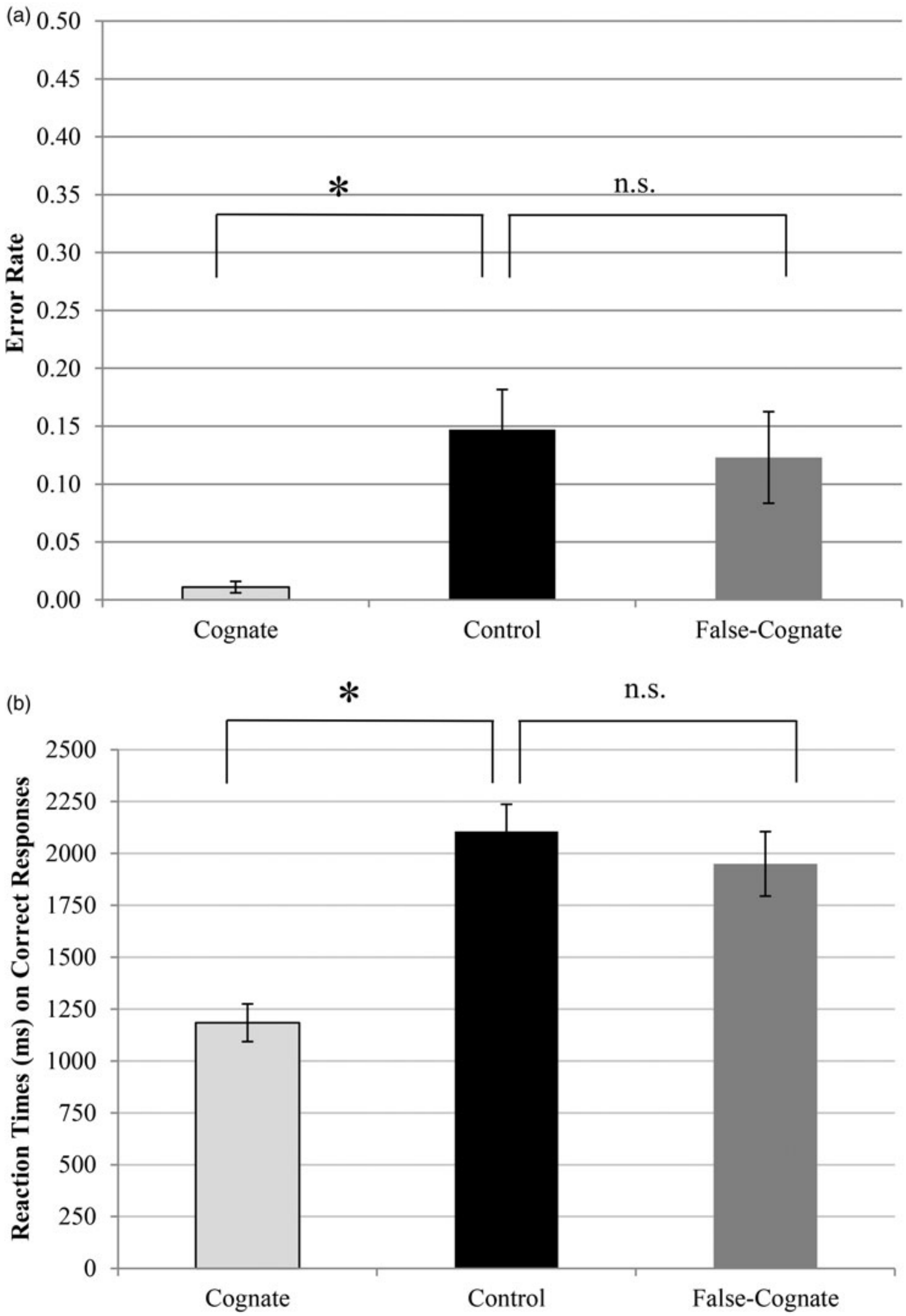

Fig. 1. The effect of Word Type on error rate (a) and RT (b) in cycle 1 (estimated means, with error bars representing SE).

from Controls. See Figure 1 and Table 4 for the results from the maximum number of participants in Cycle 1, and the Supplementary Materials (Supplementary Materials, S4) for detailed analyses of all cycles.

Performance across cycles

Due to the low number of participants who completed the 4th cycle, we focused on comparisons across the first three cycles with the 22 participants who completed these cycles. To examine modulations in performance across time and practice, the effect of Cycle (1, 2, 3, with 1 as the reference) and its interaction with Word Type were added to the model, as well as by-participant and by-item slopes for Cycle. In both the error rate and the RT data, there was a significant effect of Cycle (errors: $F(2) = 28.27$, $p < .001$; RT: $F(2{,}49.77) = 15.87$, $p < .001$), such that performance improved with Cycle (Error: $M_{Cycle\ 1} = 0.06$, $M_{Cycle\ 2} = 0.02$, $M_{Cycle\ 3} = 0.02$; RT: $M_{Cycle\ 1} = 1750$, $M_{Cycle\ 2} = 1544$, $M_{Cycle\ 3} = 1396$, see Table 5). Pairwise comparisons with Bonferroni corrections for multiple comparisons revealed that all differences among the cycles were significant, except for the error rate difference between Cycle 2 and Cycle 3. Further, there was a significant effect of Word Type (Error: $F(2) = 17.65$, $p < .001$; RT: $F(2{,}60.73) = 34.69$, $p < .001$), such that Cognates were recognized with fewer errors and more quickly than Controls, whereas False-Cognates did not differ from Controls (Error: $M_{Control} = 0.06$, $M_{Cognate} = 0.01$, $M_{False\text{-}Cognates} = 0.05$; RT: $M_{Control} = 1863$, $M_{Cognate} = 1075$, $M_{False\text{-}Cognates} = 1737$).

https://doi.org/10.1017/S1366728921000857 Published online by Cambridge University Press

Table 4. Summary of the Translation Recognition test as a function of Word Type in Cycle 1 (n=54). See Supplementary Materials (S4) for the results of other cycles.

	Error Rate				Log RT				
Fixed effect	*b*	*SE*	*z*	*P(z)*	*b*	*SE*	*df*	*t*	*P(t)*
Intercept	−1.76	0.28	−6.36	<.001***	7.65	0.06	97.34	125.09	<.001***
Word Type (Cognate)	−2.75	0.49	−5.65	<.001***	−0.58	0.09	58.42	−6.74	<.001***
Word Type (FC)	−0.21	0.41	−0.51	0.61	−0.08	0.09	57.38	−0.90	0.37
Random effect	*Variance*		*SD*		*Variance*	*SD*		*Corr.*	
Item (intercept)	1.38		1.17		0.06	0.25			
Participant (intercept)	0.99		0.99		0.07	0.27			
Word Type (cognate)	-		-		0.01	0.12		−0.53	
Word Type (FC)	-		-		0.01	0.08		−0.11	0.00
Residual	-		-		0.17	0.41			
AIC			2115.7				2928.8		
Intraclass Correlation (Item)			0.24				0.21		
Intraclass Correlation (Participant)			0.17				0.25		

Note: ± $p < 0.1$ * $p < .05$; ** $p < 0.01$; *** $p < 0.001$. FC refers to false-cognates.

Critically, the effects did not interact[4]. Thus, although performance significantly improved from cycle to cycle, the difference between Cognates and Controls, or between False-Cognates and Controls, did not vary by learning cycle.

Individual differences

We examined whether overall FL vocabulary learning was modulated by individual differences and further, whether the effects of semantic-form overlap were modulated by these individual differences. We focused on three key predictors: Phonological Short-Term Memory, L1 Verbal Fluency, and L1 Use ratings. Because these variables were not significantly correlated with each other (see Table 6), they were jointly entered into the analyses, after they have been normalized.

Learning to criterion

We examined the correlation between each of the three individual differences measures and the time it took participants to reach the learning criteria. As can be seen in Table 6, overall performance was correlated with phonological short-term memory and L1 verbal fluency, such that participants with increased phonological short-term memory and L1 verbal fluency reached the learning criterion with fewer cycles. Interestingly, performance on the False-Cognates was specifically modulated by phonological short-term memory, such that individuals with increased phonological short-term memory reached the learning criterion for these items with fewer cycles.

Performance by cycle

For cycles with enough participants (1 through 3), in each measure (error rate and RT) we specified a model including the effects of the three individual difference predictors and their interactions with Cycle and with Word Type. A maximal model including these interactions as well as by-participant and by-item intercepts, and a by-participant slope for Word Type and by-participant and by-item slopes for Cycle, was submitted to a *buildmer* function in the *buildmer* package (v. 1.3, Voeten, 2019) in R (version 3.6.1, R Core Team, 2019). With this function, the contribution of each fixed effect to the model fit is examined in a 'leave-one-out' backward procedure, such that the model is compared to a model without one of the fixed effects via a chi-squared test on the residual sum of squares of each model (similar to an *anova* function in R for model comparisons).[5] See Supplementary Materials (Supplementary Materials, S5) for full details of these selected models from the *summary* function (note that Control items are set as the reference for the Word Type effect), and Supplementary Materials (Supplementary Materials, S3) for details from the *anova* function.

Cycle 1

In the error rate analysis, phonological short-term memory predicted performance ($F(1) = 16.09$, $p < .001$) such that individuals with higher phonological short-term memory learned better

[4]To verify that the backward stepwise selection procedure we used did not inflate other terms (see Sonderegger, Wagner & Torreira, 2018) we repeated the analysis while keeping the Word Type by Cycle interaction in the model. These analyses again showed no significant interaction (Errors: $F(4) = 1.69$, $p = 0.15$; RT: $F(4,58.38) = 0.38$, $p = 0.82$) with no change in the pattern of significant effects of Word Type and Cycle.

[5]Using the forward selection procedure, instead of the backward selection direction for evaluating the contribution of each fixed effect, resulted in the exact same pattern of significant effects, with one exception. In the error data analysis of Cycle 2, the main effect of L1 verbal fluency reached significance ($F(1) = 4.73$, $p = 0.03$).

Further, including all predictors in the model, rather than removing terms (in a forward or backward procedure), resulted in the same pattern of significant effects, with two exceptions. First, in the error rate of Cycle 1, the difference between false-cognates and controls was modulated by Phonological short-term memory ($b = -0.27$, $SE = 0.14$, $z = -1.97$, $p = 0.049$), such that the difference was larger for individuals with lower phonological short-term memory ($M_{Control} = 0.29$ vs $M_{False\text{-}Cognates} = 0.36$ for individuals with lower phonological short-term memory ($Z = -2$), but $M_{Control} = 0.08$ vs $M_{False\text{-}Cognates} = 0.04$ for individuals with higher phonological short-term memory ($Z = 2$)). This reduced difference with increased phonological short-term memory likely reflects near ceiling performance in error rates for individuals with higher phonological short-term memory. Second, in the error rate analysis of Cycle 2, the effect of L1 verbal fluency reached significance ($F(1) = 5.60$, $p = 0.02$). Critically, the modulation of the difference between Cognates and Controls by phonological short-term memory remained significant ($b = 1.25$, $SE = 0.61$, $z = 2.05$, $p = 0.041$).

Table 5. Summary of the Translation Recognition test as a function of Word Type and Cycle.

	Error Rate				Log RT					
Fixed effect	*b*	*SE*	*z*	*P(z)*	*b*	*SE*	*df*	*t*	*P(t)*	
Intercept	−1.79	0.26	−6.78	<.001***	7.65	0.06	113.97	132.92	<.001***	
Word Type (Cognate)	−2.40	0.39	−6.15	<.001***	−0.56	0.07	61.32	−8.18	<.001***	
Word Type (FC)	−0.35	0.34	−1.03	0.30	−0.07	0.07	58.87	−1.09	0.28	
Cycle (2)	−1.18	0.24	−4.89	<.001***	−0.13	0.03	48.25	−4.88	<.001***	
Cycle (3)	−1.30	0.37	−3.50	<.001***	−0.22	0.05	54.95	−4.63	<.001***	
Random effect	*Var.*	*SD*	*Corr.*		*Var.*	*SD*	*Corr.*			
Item (intercept)	1.38	1.18			0.06	0.25				
Cycle 2	0.47	0.69	−0.24		0.01	0.09	−0.45			
Cycle 3	0.71	0.85	−0.64	0.55	0.06	0.25	−0.61	0.48		
Participant (intercept)	1.00	1.00			0.07	0.27				
Cycle 2	0.35	0.59	0.84		0.01	0.11	0.02			
Cycle 3	0.35	0.59	0.51	0.71	0.03	0.16	−0.09	0.54		
Type (Cognate)	-	-	-		0.01	0.11	−0.57	−0.31	−0.34	
Type (FC)	-	-	-	-	0.01	0.08	−0.22	−0.07	0.15	−0.09
Residual	-	-	-	-	0.17	0.41				
AIC			4109.8				6188.0			
Intraclass Correlation (Item)			0.22				0.21			
Intraclass Correlation (Participant)			0.16				0.25			

Note: ± $p < 0.1$ * $p < .05$; ** $p < 0.01$; *** $p < 0.001$. Intraclass correlations may not be accurate in the presence of random slopes. The overall ICC for the error rate model (adjusted ICC) is 0.47 and for the RT model it is 0.44. FC refers to false-cognates.

Table 6. Pearson correlations among the individual difference measures and learning to criterion.

Measure	1	2	3	4	5	6	7
1 Phonological short-term memory	-						
2 L1 verbal fluency	.20	-					
3 L1 use	−.00	.14	-				
4 Overall number of cycles	−.42***	−.28*	−.06	-			
5 Cognate cycles to criterion	−.11	−.09	−.00	.29*	-		
6 Control cycles to criterion	−.20	−.03	−.11	.68***	.47***	-	
7 False-Cognates cycles to criterion	−.51***	−.22	−.13	.83***	.36***	.42***	-

Note: ± $p < 0.1$ * $p < .05$; ** $p < 0.01$; *** $p < 0.001$.

than those with lower phonological short-term memory (for Control reference items $b = -0.51$, $SE = 0.13$, $z = -3.97$, $p < .001$). There was no interaction, however, between phonological short-term memory and the relevant contrasts.

In the RT analysis, L1 verbal fluency predicted performance ($F(1,53.83) = 7.54$, $p = 0.008$, for Control items $b = -0.09$, $SE = 0.03$, $df = 53.83$, $t = -2.75$, $p = 0.008$), such that individuals with better L1 verbal fluency performed more quickly than those with lower L1 verbal fluency. However, there were no interactions between L1 verbal fluency and the critical contrasts.

Cycle 2

In the error rate analysis, individuals with higher L1 use performed with fewer errors than those with lower L1 use ($F(1) = 6.54$, $p = 0.01$, for Control items $b = -0.48$, $SE = 0.18$, $z = -2.65$, $p = 0.008$). In addition, individual differences in phonological short-term memory modulated the difference between Cognates and Controls ($b = -1.09$, $SE = 0.49$, $z = 2.19$, $p = 0.03$), such that it was larger for individuals with lower phonological short-term memory (see Figure 2). Specifically, whereas performance on Control words improved with higher phonological short-term memory, performance on Cognates remained stable, potentially due to near ceiling performance.

In the RT analysis, the difference between False-Cognates and Controls was modulated by phonological short-term memory ($b = -0.07$, $SE = 0.03$, $df = 1672$, $t = -2.57$, $p = 0.01$). Increased phonological short-term memory was associated with a larger advantage for False-Cognates over Control items (see Figure 3). Whereas RTs for False-Cognates improved with phonological short-term memory, performance on Control items remained unchanged.

https://doi.org/10.1017/S1366728921000857 Published online by Cambridge University Press

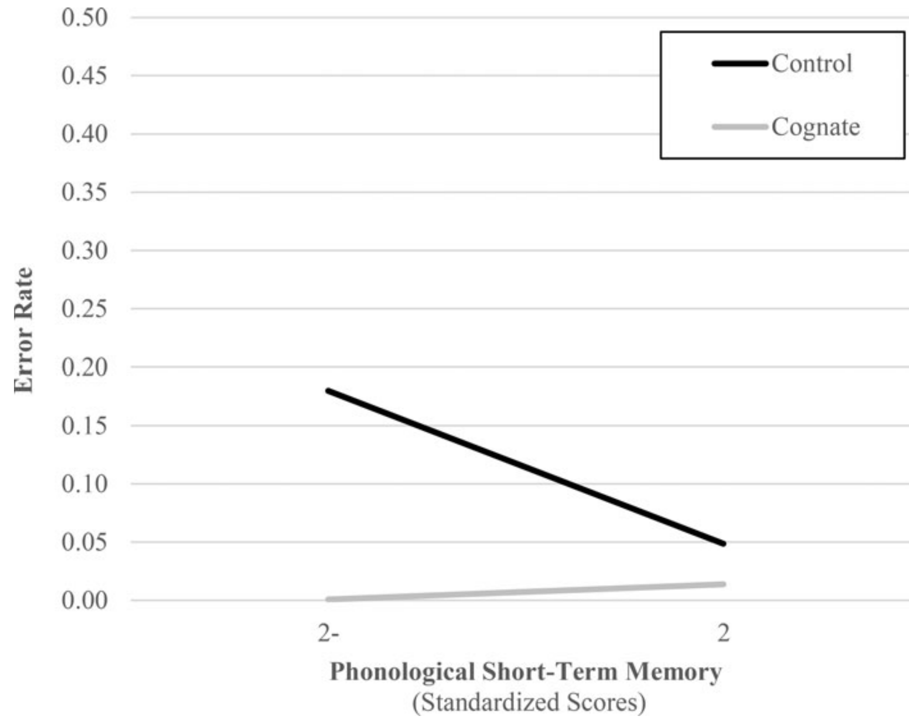

Fig. 2. Estimated difference between Cognate and Control items as a function of Phonological Short-Term memory in the error rate data of Cycle 2.

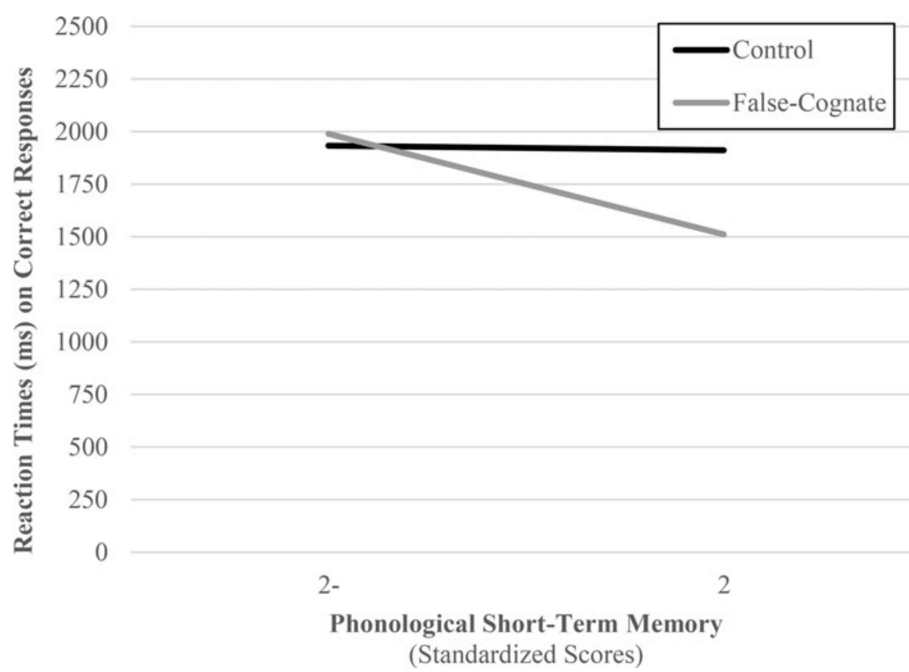

Fig. 3. Estimated difference between False-Cognates and Control items as a function of Phonological Short-Term memory in the RT data of Cycle 2.

At the same time, the difference between False-Cognates and Controls was modulated by L1 verbal fluency ($b = 0.06$, $SE = 0.02$, $df = 1672$, $t = 2.56$, $p = 0.01$), such that increased L1 verbal fluency was associated with faster processing of Control items but not of False-Cognates, such that the difference between them was reduced (see Figure 4).

Cycle 3

Phonological short-term memory modulated the error rate difference between Cognates and Controls ($b = 1.03$, $SE = 0.39$, $z = 2.62$, $p = 0.009$). In particular, as in Cycle 2, whereas individuals with lower phonological short-term memory exhibited fewer errors for Cognates relative to Controls (for individuals 1 *SD* below the mean $M_{Control} = 0.15$ vs. $M_{Cognate} = 0.01$), there was no difference for those with higher phonological short-term memory (for individuals 1 *SD* above the mean $M_{Control} = 0.09$ vs. $M_{Cognate} = 0.04$), likely due to near ceiling performance on this measure on Cognate items. The (lack of) difference between False-Cognates and Controls was not modulated by individual differences.

In the RT analysis, L1 verbal fluency emerged as a significant predictor ($F(1,21.97) = 6.85$, $p = 0.02$, for Control items $b = -0.14$, $SE = 0.05$, $df = 21.97$, $t = 2.62$, $p = 0.02$), such that increased L1

https://doi.org/10.1017/S1366728921000857 Published online by Cambridge University Press

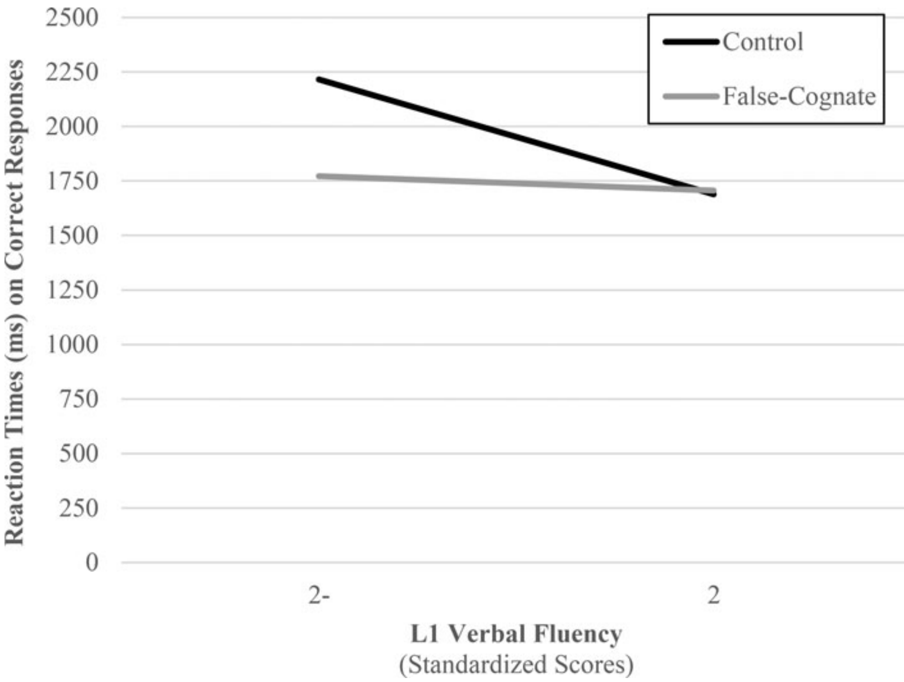

Fig. 4. Estimated difference between False-Cognates and Control items as a function of L1 Verbal Fluency in the RT data of Cycle 2.

verbal fluency was associated with shorter RTs. This factor did not modulate the effect of Word Type.

Backward influence analyses

To examine whether learning of Arabic vocabulary influenced processing of Hebrew words, indicative of backward cross-language influences, we examined learners' performance in a visual lexical decision task on Hebrew letter strings. Data from ten participants were excluded because they did not reach the 80% learning criteria in one or more of the word types; data from one additional participant were excluded because they reported incorrectly pressing the buttons at the beginning of the lexical decision task. Thus, data from 43 participants that reached the learning criterion (80% success) in all word types were analyzed for the backward influences task.

Table 7 presents mean performance in the Hebrew lexical decision task. As is typically observed in lexical decision tasks (e.g., Balota et al., 2007; Peleg et al., 2020), words were responded to significantly more quickly ($b = 0.09$, $SE = 0.01$, $df = 65.73$, $t = 6.95$, $p < .001$), and accurately ($b = 0.90$, $SE = 0.34$, $z = 2.67$, $p = 0.008$) than non-words. Of relevance, model comparisons using the *buildmer* function revealed that the effect of Word Type was not significant in the error rate data, but was significant in the RT data ($F(2,50.90) = 7.51$, $p = 0.001$), such that False-Cognate items were responded to more slowly than Control items ($b = 0.07$, $SE = 0.02$, $df = 50.95$, $t = 3.76$, $p < .001$). The difference between Cognate and Control items was not significant ($b = 0.01$, $SE = 0.02$, $df = 50.70$, $t = 0.52$, $p = 0.61$).

To verify whether the difference between False-Cognates and Controls for the learners indicate the presence of backward influences, we compared their performance to that of another group of 30 native Hebrew speakers who did not participate in the learning paradigm (see Supplementary Materials, S6 for background characteristics and comparisons to the learner group). These analyses including both groups (see Table 8) revealed no effects in the error rate data, but a main effect of Word Type in the RT data ($F(2,50.85) = 5.78$, $p = 0.005$), with faster responses to Control relative to False-Cognate items ($b = 0.02$, $SE = 0.01$, $df = 50.92$, $t = 3.12$, $p = 0.003$), and no difference between Control and Cognate items ($b = -0.00$, $SE = 0.02$, $df = 50.69$, $t = -0.11$, $p = 0.91$). The learners responded marginally[6] more slowly ($F(2,76.49) = 3.38$, $p = 0.07$), likely because this was the last task in their learning and testing protocol. Critically, the interaction between the two factors was marginally significant ($F(2,49.08) = 3.13$, $p = 0.05$), but the specific difference between False-Cognates and Control items was modulated by group ($b = 0.06$, $SE = 0.03$, $df = 49.33$, $t = 2.35$, $p = 0.023$). Pairwise comparisons with Bonferroni corrections revealed that the difference between False-Cognates and Controls was significant for the learner group (*value* $= -0.07$, $\chi^2 = 14.20$, $p < .001$), but not for the non-learner group (*value* $= -0.01$, $\chi^2 = 0.22$, $p = 1.00$, see Figure 5). Thus, learning the Arabic meaning of the phonological forms of False-Cognates to an 80% accuracy criterion resulted in backward influences in L1 processing immediately after learning. A Cognate facilitation effect, however, was not observed in L1 Hebrew.

We further examined whether within the learner group, individual differences modulated the effect. These analyses revealed that participants' overall learning score from the translation recognition test and the three individual difference predictors of interest (phonological short-term memory, L1 verbal fluency, L1 use) did not modulate the difference across word types.

[6]Because the interaction of Word Type and Group was only marginally significant, the interaction term was not maintained by the *buildmer* selection procedure, which drops fixed effects with $p > .05$. Thus, the values presented in Table 8 and used for Figure 5 are from a model in which the random structure selected by the *buildmer* ((1|Participant)+(1+Group|Item)) was adopted, while keeping the interaction term in the model. We view this as the more conservative approach, because when using the *buildmer* function with forward selection criteria, or when keeping all fixed effects in, the selected model included a significant (rather than marginally significant) interaction between Word Type and Group, with no by-item slope in the random structure.

https://doi.org/10.1017/S1366728921000857 Published online by Cambridge University Press

Table 7. Error Rate and RT on correct responses in the Hebrew lexical decision task as a function of Lexicality, Word Type and Group.

		Word Type			Lexicality	
Group	Measure	Cognate	False Cognates	Control	Words	Non-words
Learners	Error Rate	0.04 (0.01)	0.05 (0.01)	0.03 (0.00)	0.04 (0.00)	0.07 (0.01)
	RT (in ms)	646 (16)	686 (17)	627 (10)	672 (7)	834 (9)
Non-Learners	Error Rate	0.03 (0.01)	0.05 (0.01)	0.03 (0.01)	0.04 (0.00)	0.07 (0.01)
	RT (in ms)	576 (12)	605 (13)	599 (10)	602 (6)	717 (7)

Note: Standard errors appear in parentheses.

Table 8. Summary of the Lexical Decision task as a function of Word Type and Group.

	Error Rate				Log RT				
Fixed effect	*b*	*SE*	*z*	*P(z)*	*b*	*SE*	*df*	*t*	*P(t)*
Intercept	−4.73	0.49	−9.60	<.001***	6.35	0.03	99.17	206.72	<.001***
Group (Learners)	−0.90	0.64	−1.41	0.16	0.04	0.04	87.47	1.02	0.31
Word Type (Cognate)	−0.07	0.44	−0.17	0.86	−0.03	0.03	47.90	−1.27	0.21
Word Type (FC)	0.46	0.40	1.15	0.25	0.01	0.03	48.40	0.47	0.64
Group*Word Type (Cognate)	0.76	0.59	1.29	0.20	0.04	0.03	48.72	1.65	0.10
Group*Word Type (FC)	0.53	0.55	0.97	0.33	0.06	0.03	49.33	2.35	0.02*
Random effect	*Variance*		*SD*		*Variance*		*SD*	*Corr.*	
Item (intercept)	0.39		0.55		0.00		0.06		
Group (Learners)	-		-		0.00		0.05	−0.72	
Participant (intercept)	3.34		1.83		0.02		0.14		
Residual	-		-		0.07		0.26		
AIC			767.2				909.8		
Intraclass Correlation (Item)			0.04				0.04		
Intraclass Correlation (Participant)			0.48				0.22		

Note: ± $p < 0.1$ * $p < .05$; ** $p < 0.01$; *** $p < 0.001$. Intraclass correlations may not be accurate in the presence of random slopes. The overall ICC for the RT model (adjusted ICC) is 0.25. FC refers to false-cognates.

Discussion

In the current study, we investigated FL vocabulary learning, focusing on learning FL words which overlap in form and meaning (cognates), or in form only (false-cognates), with learners' L1. Furthermore, we examined whether individual differences in phonological short-term memory, L1 verbal fluency and L1 use modulate the learning process across the different word types. Finally, we tested for backward influences on learners' L1 following the learning of new meanings to existing forms (false-cognates). The results of the study show that cognates were learned more quickly, such that participants reached the 80% learning criterion with the fewest number of learning and testing cycles. Cognates were also processed more quickly and accurately than controls in the recognition test. The cognate advantage was more pronounced for individuals with lower phonological short-term memory abilities. Critically, whereas overall performance suggests no difference in learning false-cognates and control items, the individual difference analyses reveal that individuals with higher phonological short-term memory abilities and those with lower L1 verbal fluency abilities exhibited a significant advantage for false-cognates over control items. Finally, there was evidence for backward influences of learning false-cognates, such that false-cognates were responded to more slowly than control words following learning. We discuss each of these findings and their implications in the following sections.

The cognate advantage in learning

In concurrence with previous studies, cognate words were easier and faster to learn compared to control words (e.g., Lotto & De Groot, 1998). As expected, participants reached the learning criterion for cognates with the fewest number of learning and testing cycles and performed more quickly and more accurately in the recognition test on cognate items. Of relevance, whereas previous studies typically focused on cognate learning between languages that share orthography (Lotto & De Groot, 1998; De Groot & Keijzer, 2000; Raboyeau et al., 2010), the current study demonstrated the cognate learning advantage when learning a FL via phonology, with no reliance on a shared orthographic system (see also Ghazi-Saidi & Ansaldo, 2017). As such, the findings

https://doi.org/10.1017/S1366728921000857 Published online by Cambridge University Press

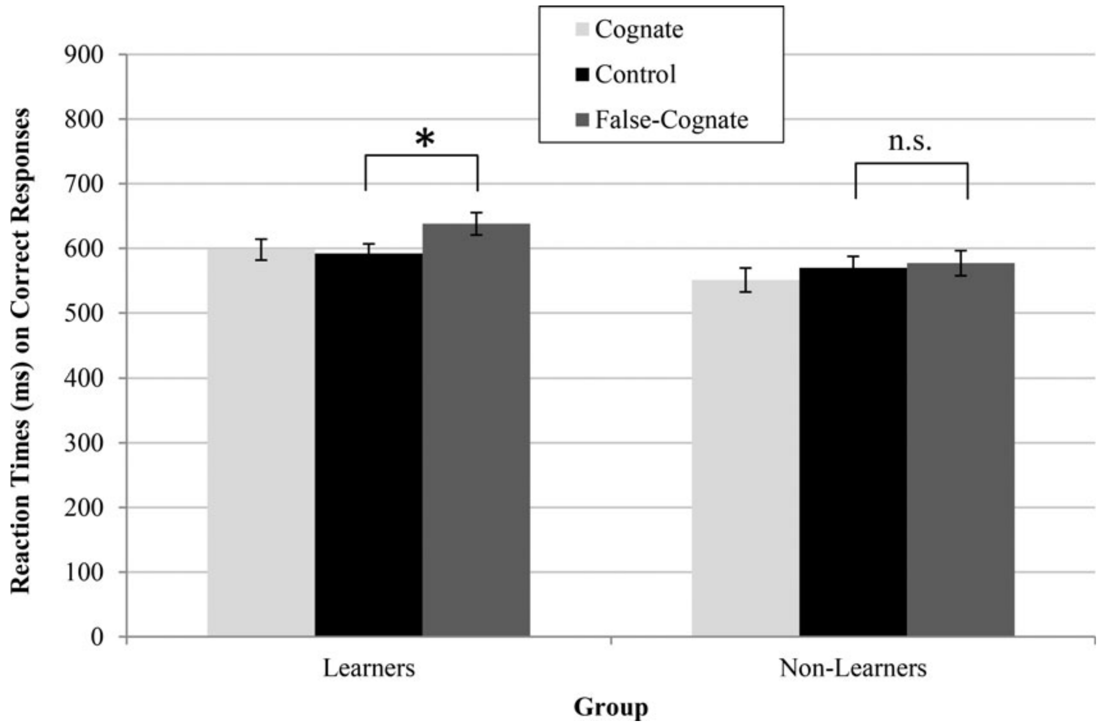

Fig. 5. Estimated difference between False Cognates and Control items in the RT data of the Lexical Decision Task as a function of Group.

underscore the generalizability of the effect and highlight the role of phonological overlap in mediating this advantage. Moreover, although the cognates used in the current study substantially overlap in form across Hebrew and Arabic, their phonological realization is not identical. In fact, none of the 14 cognates used would be considered an identical cognate (see Supplementary Materials, S1). Nonetheless, cognate words were learned substantially better than non-cognates.

By examining individual difference modulations of FL learning we observed a greater cognate advantage for individuals with lower phonological short-term memory. However, we believe this is because performance on cognate items at this point was near ceiling for all participants. As seen in Figure 2, increased phonological short-term memory was associated with fewer errors for control items, but no such facilitation was observed for cognate items because performance on these easier items was already at ceiling.

Individual differences in learning of false cognates

The more critical aim of the current study was to test whether false-cognates are easier or harder to learn compared to control words. The results show no overall difference between learning false-cognates and control words, as revealed by the number of cycles required to meet the learning criterion, or by the latency and error rate data in the recognition tests.

The comparable overall performance on false-cognates and control words may reflect the operation of two opposing mechanisms, a form facilitation and a meaning competition. Specifically, the overlap in form across the L1 and the to-be-learned words may facilitate learning compared to control items (Mulík et al., 2019; Marecka, Szewczyk, Otwinowska, Durlik, Foryś-Nogala, Kutyłowska & Wodniecka, 2020). At the same time, because the meanings of false-cognates differ from those of the L1, meaning competition for these items may cause difficulty in learning false-cognates (Rodd et al., 2012). These two mechanisms may both be at play and may cancel each other out, leading to an overall null effect. Indeed, Ghazi-Saidi and Ansaldo (2017) similarly observed no learning accuracy or RT difference between false-cognates and controls, but found differential brain recruitment when processing the two types of items. Thus, the comparable performance on false-cognates and control words may not be supported by the same learning mechanism. Furthermore, Fang et al. (2017) exemplified differences in the time course of these opposing mechanisms within a language, such that initial form facilitation was followed by subsequent meaning interference. It is possible that the current study taps a point in time at which the two processes balance out. This claim is further supported by the individual differences results.

Specifically, we observed that processing of false-cognates was modulated by phonological short-term memory. First, individuals with higher phonological short-term memory required fewer learning cycles to learn false-cognates compared to those with lower phonological short-term memory. Second, individuals with higher phonological short-term memory were able to recognize false-cognate words significantly more quickly than control items in the 2nd cycle of learning. This finding can be explained by considering the suggestion that learning of false-cognates is affected both by facilitation due to the overlap in phonological form and by competition between the different semantic representations (Fang et al., 2017). We propose that individuals with higher phonological short-term memory relied to a greater extent on the phonological form, and thus exhibited an advantage for false-cognates over control words.

At the same time, we observed learning modulations with L1 verbal fluency. Specifically, increased L1 verbal fluency was

https://doi.org/10.1017/S1366728921000857 Published online by Cambridge University Press

expected to be associated with better learning, as seen in Figure 4 for control items. Critically, this improved learning was not observed for false-cognates. We propose that learning of false-cognates is not facilitated by increased verbal fluency because of increased meaning competition for individuals with higher verbal fluency. This sensitivity to competition for individuals with higher verbal fluency is consistent with the findings of Degani and Goldberg (2019, Figure 4), where increased verbal fluency was associated with better learning of unambiguous (control) items, but not of ambiguous items. As a result, in that study the ambiguity disadvantage in learning was more pronounced for individuals with higher verbal fluency. Here, we observed that individuals with lower verbal abilities exhibited a false-cognate advantage over control words, as they were able to benefit from form facilitation, and did not rely on meaning enough to suffer from competition. Individuals with higher verbal abilities did not show a difference because for them, meaning competition offset the form facilitation.

This finding is important for two reasons. First, it shows that the lack of overall difference in learning false-cognates and controls likely obscures a substantial difference in the processes involved in learning these different types of items (Ghazi-Saidi & Ansaldo, 2017). Learners with different cognitive and linguistic resources may be differentially affected by the overlap in form and not in meaning. Second, the finding underscores the importance of considering individual differences in learners' prior knowledge and abilities and shows the merits in jointly considering item and learner characteristics (Degani & Goldberg, 2019).

Our findings demonstrate that the advantage for learning false-cognates over control words may become apparent under some conditions, but not others, but critically does not reverse into a false-cognate disadvantage in learning. Converging evidence comes from a recently published paper by Marecka et al. (2020), who examined learning of Polish non-words in association with pictures. These non-words could resemble the Polish word for the associated picture (i.e., cognates), resemble a different Polish word (i.e., false-cognates) or be different from existing Polish words (i.e., non-cognate controls). Of relevance, their findings show no difference in learning false-cognates versus non-cognates in a recognition task, but a false-cognate advantage in a production task. The authors argue that the interference with L1 meaning was either nonexistent or outweighed by the form overlap facilitation, when production of the form was examined.

Individual differences in FL learning

Furthermore, the current findings show an important overall relation between FL vocabulary learning on the one hand, and phonological short-term memory and L1 verbal fluency on the other. Consistent with previous work (Degani & Goldberg, 2019), we found that phonological short-term memory was positively correlated with overall learning, such that individuals with higher phonological short-term memory required fewer learning cycles to reach the learning criterion and performed the translation recognition tests more accurately across cycles.

We further found that participants' verbal fluency in their L1 was positively linked to learning, as reflected by both the number of learning cycles to criterion, and the translation recognition RT data across cycles. This finding is consistent with previous research where verbal fluency in the language through which learning took place was positively linked to learning (Degani & Goldberg, 2019). Notably, however, the verbal fluency measure may not capture only linguistic abilities, but also domain general abilities such as executive functions (Jurado & Rosselli, 2007; Friesen et al., 2015). Indeed, in the current study, we observed only low correlations between the L1 verbal fluency score and the subjective L1 and L2 proficiency and use measures (see also Degani & Goldberg, 2019), but this may be due to limitations of the self-report measure (Tomoschuk et al., 2019). Future studies in which additional measures of executive control abilities are collected would shed light on the mechanisms underlying the relation between verbal fluency and FL vocabulary learning. Moreover, executive functions may be of special relevance in the case of false-cognate learning. In particular, false-cognates require the management of meaning competition between the two languages. Therefore, learning and processing of these items may recruit executive control to a greater degree than control words. Indeed, Grant et al. (2015) tested for the role of cognitive control abilities amongst English speakers learning false-cognates (interlingual homographs) and unambiguous Spanish (L2) words throughout an academic year. The findings showed that inhibitory control was more critical in early rather than advanced stages of L2 learning (consistent with this, Prior et al., 2017 did not observe executive control modulations of false-cognate processing among proficient Arabic–Hebrew bilinguals). Thus, future studies in which individual differences in executive control are collected may be of special interest in the case of false-cognate learning.

To the extent that L1 verbal fluency is taken to reflect linguistic abilities, the observed relation between verbal fluency and learning is informative with respect to the debate over whether L1 verbal fluency is linked to FL learning due to a common proficiency construct, or due to direct transfer. If L1 verbal fluency reflects a common proficiency construct (e.g., Cummins, 1979, 1991; for a cognitive variant, see Geva, 2014), the difference between individuals with lower and higher abilities was predicted to be larger under more demanding learning conditions: namely, false-cognates and controls, but not cognates. The data shows that although L1 verbal fluency did modulate the difference between false-cognates and controls, it was mostly the control items that were learned better with increased L1 verbal fluency. According to transfer accounts, by which the advantage of individuals with higher L1 verbal fluency stems from their specific reliance on L1 phonological and semantic representations in their lexicon (Gathercole, 2006), we expected individuals with higher L1 verbal fluency to show a larger advantage for cognates because these items overlap in both form and meaning across languages. We further expected these individuals to show a decrement in processing false-cognates because of the misalignment in meaning across languages. The findings support the second of these predictions, but not the first, because processing of cognates does not appear to vary by L1 verbal fluency in the current study. Our results highlight the possibility that both aspects are in fact at play: prior language abilities may facilitate FL learning both because they allow for direct transfer, and because they indirectly reflect a more general enhanced language ability of the individual (for further discussion see e.g., Hirosh & Degani, 2018).

Finally, the current findings exemplify not only that FL learning is modulated by individual differences, but also that these individual differences may differentially affect learning of different types of words. Nonetheless, limitations in power may have obscured additional variability of interest. Thus, future studies incorporating larger samples may prove useful in examining such interactions, especially because the number of items that learners can successfully learn may be limited in single session learning protocols of the type implemented here.

Backward influence

The third and final aim of this study was to test whether learning new meanings to already established forms (false-cognates) would result in backward influences, such that the newly learned meanings would affect processing of the word form in the L1. Previous findings demonstrated backward influences as a result of altering L1 mappings following novel word learning within a language (Davis, Di Betta, Macdonald & Gaskell, 2009; Fang & Perfetti, 2019; Malt et al., 2016; Rodd et al., 2012; Maciejewski et al., 2020), and thus, we expected learners to respond to false-cognates as ambiguous Hebrew words entailing more than one meaning after having learned new meanings to these words. At the same time, we expected cognates to enjoy an advantage in processing due to the convergence of meanings from both languages. The results of the current study provide initial evidence for such backward influences, in that lexical decision latencies to false-cognate Hebrew words were slower immediately after learning compared to control items. This increased processing time was not present for Hebrew speakers who did not learn the Arabic vocabulary. Cognate words, in contrast, were processed similarly by the two groups, suggesting no cognate facilitation immediately following learning.

Successful learning of false-cognate words to an 80% criterion was sufficient to lead to a false-cognate interference in the L1. Thus, having learned a competing meaning in a FL to the same phonological form made it slightly more difficult for native Hebrew speakers to deem these words as Hebrew words in a lexical decision task. Notably, the current study may have underestimated the presence of backward influences because these were examined immediately following learning. Previous work suggests that knowledge consolidation may be necessary for backward influences of the type examined here to appear. For instance, in a study conducted by Davis et al. (2009), adult participants learned non-words as novel words (e.g., alcohin) and completed a lexical decision task on the day of learning and on the following day. Results showed competition between these novel non-words and similar-sounding established English words (e.g., alcohol) only in the day following learning, but not on the same day as learning. Similarly, Dumay and Gaskell (2007) demonstrated that a consolidation period including a night sleep is necessary in order to observe lexical competition effects of newly learned spoken forms (see also Dumay & Gaskell, 2012). Moreover, consolidation in the L2 has been suggested to require more time than consolidation of L1 novel words (Lindsay & Gaskell, 2010). Future studies in which participants will be given the opportunity for overnight consolidation will test the possibility that false-cognates and cognates learned in a FL can similarly impact processing of L1 words to a greater extent after consolidation.

Conclusion

The current study shows that the overlap of form and meaning across the to-be-learned FL vocabulary and existing linguistic knowledge impacts vocabulary learning, and critically that these effects vary with learner characteristics. While all learners exhibited an advantage in learning cognates, these effects were larger for individuals with lower phonological short-term memory abilities, who were further from ceiling performance. Further, while there was no overall difference in learning false-cognates vs. control words, likely reflecting the operation of opposing mechanisms of form facilitation and meaning competition, individuals with higher phonological short-term memory and those with lower L1 verbal fluency exhibited a learning advantage for false-cognates over control words. Hence, the current findings carry important implications for educational settings. Specifically, although this study focused on learning in laboratory rather than natural settings (see Tokowicz & Degani, 2015 for discussion), in light of the variability in the learning process observed here, vocabulary learning in educational settings should aim to take into account each individual's linguistic and cognitive abilities in order to provide them with optimal learning opportunities. Further, the importance of considering each word type separately is highlighted by recent work showing that some instructional conditions (e.g., language of instruction – namely, learning via the L1 or the L2 among bilinguals learning a FL) affects learning of false cognates and control words, but not cognate words (Hirosh & Degani, 2021). Together, these findings highlight the relevance of jointly taking into account item-based and learner-based characteristics and differences in both linguistic and cognitive abilities in the early stages of FL learning. Finally, the study provides initial novel evidence for backward influences, in that processing of L1 words was affected immediately after learning competing FL meanings to L1 forms, although no substantial consolidation phase has taken place. The findings thus highlight the importance of cross-language mapping in the processes involved in vocabulary learning and the emergence of the bilingual lexicon.

Acknowledgements. This work was funded by the Language Learning Small Grants Research Program and by EU-FP7 grant CIG-322016 to Tamar Degani.

Supplementary Material. For supplementary material accompanying this paper, visit https://doi.org/10.1017/S1366728921000857

References

Baayen RH, Davidson DJ and Bates DM (2008) Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* **59**, 390–412. doi:10.1016/j.jml.2007.12.005

Balota DA, Yap MJ, Cortese MJ, Hutchison KA, Kessler B, Loftis B, Neely JH, Nelson DL, Simpson GB and Treiman R (2007) The English Lexicon Project. *Behavior Research Methods* **39**(3), 445–459.

Bartolotti J and Marian V (2012) Language learning and control in monolinguals and bilinguals. *Cognitive Science* **36**(6), 1129–1147. doi:10.1111/j.1551-6709.2012.01243.x

Bates D, Kliegl R, Vasishth S and Baayen H (2015) Parsimonious mixed models. *arXiv preprint arXiv*:1506.04967.

Bates D, Maechler M, Bolker B and Walker S (2015) Fitting Linear Mixed–Effects Models Using lme4. *Journal of Statistical Software* **67**, 1–48. doi:10.18637/jss.v067.i01

Bell A, Fairbrother M and Jones K (2019) Fixed and random effects models: making an informed choice. *Quality & Quantity* **53**(2), 1051–1074. doi:10.1007/s11135-018-0802-x

Bracken J, Degani T, Eddington C and Tokowicz N (2017) Translation semantic variability: How semantic relatedness affects learning of translation-ambiguous words. *Bilingualism: Language and Cognition* **20**(4), 783. doi:10.1017/S1366728916000274

Brauer M and Curtin JJ (2018) Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods* **23**(3), 389–411. doi:10.1037/met0000159

Butler AC, Karpicke JD and Roediger III, HL (2008) Correcting a metacognitive error: feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **34**(4), 918. doi:10.1037/0278-7393.34.4.918

Carpenter E, Rao L, Peñaloza C and Kiran S (2020) Verbal fluency as a measure of lexical access and cognitive control in bilingual persons with aphasia. *Aphasiology* **34** (11), 1341–1362. doi:10.1080/02687038.2020.1759774

Cheung H (1996) Nonword span as a unique predictor of second-language vocabulary language. *Developmental Psychology* **32**(5), 867. doi:10.1037/0012-1649.32.5.867

Cummins J (1979) Linguistic interdependence and educational development of bilingual children. *Review of Educational Research* **49**(2), 222–251. doi:10.3102/00346543049002222

Cummins J (1991) Interdependence of first-and second-language proficiency in bilingual children. In Bialystok E (Ed.), *Language Processing in Bilingual Children*. Cambridge: Cambridge University Press, pp. 70–89. doi: 10.1017/CBO9780511620652.006

Davis MH, Di Betta AM, Macdonald MJ and Gaskell MG (2009) Learning and consolidation of novel spoken words. *Journal of Cognitive Neuroscience* **21**(4), 803–820. doi:10.1162/jocn.2009.21059

Degani T and Tokowicz N (2010) Ambiguous words are harder to learn. *Bilingualism: Language and Cognition* **13**(3), 299–314. doi:10.1017/S1366728909990411

Degani T, Prior A and Tokowicz N (2011) Bidirectional transfer: The effect of sharing a translation. *Journal of Cognitive Psychology* **23**(1), 18–28. doi:10.1080/20445911.2011.445986

Degani T, Tseng AM and Tokowicz N (2014) Together or apart: Learning of translation-ambiguous words. *Bilingualism: Language and Cognition* **17**(4), 749–765. doi:10.1017/S1366728913000837

Degani T, Prior A and Hajajra W (2018) Cross-language semantic influences in different script bilinguals. *Bilingualism: Language and Cognition* **21**(4), 782–804. doi:10.1017/S1366728917000311

Degani T and Goldberg M (2019) How Individual Differences Affect Learning of Translation-Ambiguous Vocabulary. *Language Learning* **69** (3), 600–651. doi:10.1111/lang.12344

De Groot AM and 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. doi:10.1111/0023-8333.00110

Dijkstra T and van Heuven WJB (2002) The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition* **5**(3), 175–197. doi:10.1017/S1366728902003012

Dumay N and Gaskell MG (2007) Sleep-associated changes in the mental representation of spoken words. *Psychological Science* **18**(1), 35–39. doi: org/10.1111/j.1467-9280.2007.01845.x

Dumay N and Gaskell MG (2012) Overnight lexical consolidation revealed by speech segmentation. *Cognition* **123**(1), 119–132. doi:10.1016/j.cognition.2011.12.009

Ecke P (2015) Parasitic vocabulary acquisition, cross-linguistic influence, and lexical retrieval in multilinguals. *Bilingualism: Language and Cognition* **18** (2), 145–162. doi:10.1017/S1366728913000722

Fang X, Perfetti C and Stafura J (2017) Learning new meanings for known words: biphasic effects of prior knowledge. *Language, Cognition and Neuroscience* **32**(5), 637–649. doi:10.1080/23273798.2016.1252050

Fang X and Perfetti CA (2019) Learning new meanings for known words: Perturbation of original meanings and retention of new meanings. *Memory & Cognition* **47**(1), 130–144. doi:10.3758/s13421-018-0855-z

Floyd S and Goldberg AE (2020) Children make use of relationships across meanings in word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. doi:10.1037/xlm0000821

Friesen DC, Luo L, Luk G and Bialystok E (2015) Proficiency and control in verbal fluency performance across the lifespan for monolinguals and bilinguals. *Language, Cognition and Neuroscience* **30** (3), 238–250. doi:10.1080/23273798.2014.918630.

Gathercole SE (2006) Nonword repetition and word learning: The nature of the relationship. *Applied psycholinguistics* **27**(4), 513–543. doi:10.1017/S0142716406060383

Gathercole SE, Hitch GJ and Martin AJ (1997) Phonological short-term memory and new word learning in children. *Developmental psychology* **33** (6), 966–979. doi:10.1037/0012-1649.33.6.966

Geva E (2014) The cross-language transfer journey—a guide to the perplexed. *Written Language and Literacy* **17**(1), 1–15. doi:10.1075/wll.17.1.01gev

Ghazi-Saidi L and Ansaldo AI (2017) The neural correlates of semantic and phonological transfer effects: language distance matters. *Bilingualism: Language and Cognition* **20**(5), 1080–1094. doi:10.1017/S136672891600064X

Gollan TH, Montoya RI and Werner GA (2002) Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology* **16**(4), 562–576. doi:10.1037/0894-4105.16.4.562

Grant AM, Fang SY and Li P (2015) Second language lexical development and cognitive control: A longitudinal fMRI study. *Brain and language* **144**, 35–47. https://doi.org/10.1016/j.bandl.2015.03.010

Hirosh Z and Degani T (2018) Direct and indirect effects of multilingualism on novel language learning: An integrative review. *Psychonomic Bulletin & Review* **25**(3), 892–916. doi:10.3758/s13423-017-1315-7

Hirosh Z and Degani T (2021) Novel Word Learning Among Bilinguals Can Be Better Through the (Dominant) First Language Than Through the Second Language. *Language Learning*. doi:10.1111/lang.12457

Hulstijn JH (2011) Language proficiency in native and nonnative speakers: An agenda for research and suggestions for second-language assessment, *Language Assessment Quarterly* **8**(3), 229–249. doi:10.1080/15434303.2011.565844

Ibrahim R (2006) Morpho-phonemic similarity within and between languages: A factor to be considered in processing Arabic and Hebrew. *Reading and Writing* **19**(6), 563–586. doi:10.1007/s11145-006-9009-y

Johns MA and Steuck J (2021) Is codeswitching easy or difficult? Testing processing cost through the prosodic structure of bilingual speech. *Cognition* **211**, 104634. doi:10.1016/j.cognition.2021.104634

Jurado MB and Rosselli M (2007) The elusive nature of executive functions: a review of our current understanding. *Neuropsychology Review* **17**(3), 213–233. doi:10.1007/s11065-007-9040-z

Kang SH, Gollan TH and Pashler H (2013) Don't just repeat after me: Retrieval practice is better than imitation for foreign vocabulary learning. *Psychonomic Bulletin & Review* **20**(6), 1259–1265. doi:10.3758/s13423-013-0450-z

Karpicke JD (2012) Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science* **21**(3), 157–163. doi:10.1177/0963721412443552

Karpicke JD and Roediger III, HL (2008) The critical importance of retrieval for learning. *Science* **319**, 966–968. doi:10.1126/science.1152408

Kaushanskaya M (2012) Cognitive mechanisms of word learning in bilingual and monolingual adults: The role of phonological memory. *Bilingualism: Language and Cognition* **15**(3), 470–489. doi:10.1017/S1366728911000472

Kavé G (2005) Phonemic fluency, semantic fluency, and difference scores: Normative data for adult Hebrew speakers. *Journal of Clinical and Experimental Neuropsychology* **27**(6), 690–699. doi:10.1080/13803390490918499

Keuleers E and Brysbaert M (2010) Wuggy: A multilingual pseudoword generator. *Behavior research methods* **42**(3), 627–633. doi:10.3758/BRM.42.3.627

Kilgarriff A, Baisa V, Bušta J, Jakubíček M, Kovář V, Michelfeit J, Rychlý P and Suchomel V (2014) The sketch engine: Ten years on. *Lexicography* **1** (1), 7–36. doi:10.1007/s40607-014-0009-9.

Kilgarriff A, Reddy S, Pomikálek J and Avinesh PVS (2010) A corpus factory for many languages. In *LREC workshop on web services and processing pipelines*, Malta, May 2010 (pp. 2010).

Kroll JF and 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., Bobb SC and Wodniecka Z (2006) Language selectivity is the exception, not the rule: Arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition* **9**(2), 119–135. doi:10.1017/S1366728906002483

Kuznetsova A, Brockhoff PB and Christensen RHB (2017) lmer test package: Tests in linear mixed effects models. *Journal of Statistical Software* **82**, 1–26. doi:10.18637/jss.v082.i13

https://doi.org/10.1017/S1366728921000857 Published online by Cambridge University Press

Lenth R (2020) emmeans: Estimated marginal means, aka least-squares means. R package version 1.5.2-1. Retrieved from https://CRAN.R-project.org/package=emmeans

Linck JA, Kroll JF and Sunderman G (2009) Losing access to the native language while immersed in a second language: Evidence for the role of inhibition in second-language learning. *Psychological science* **20**(12), 1507–1515. doi: org/10.1111/j.1467-9280.2009.02480.x

Lindsay S and Gaskell MG (2010) A complementary systems account of word learning in L1 and L2. *Language Learning* **60**, 45–63. doi:10.1111/j.1467-9922.2010.00600.x

Lotto L and de Groot AMB (1998) Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning* **48**, 31–69. doi:10.1111/1467-9922.00032

Lüdecke D, Makowski D, Waggoner P and Patil I (2020) performance: Assessment of Regression Models Performance. R Package, Retrieved from https://easystats.github.io/performance/

Maciejewski G, Rodd JM, Mon-Williams M and Klepousniotou E (2020) The cost of learning new meanings for familiar words, *Language, Cognition and Neuroscience* **35**(2), 188–210. doi:10.1080/23273798.2019.1642500

MacWhinney B (2005) A unified model of language acquisition. In Kroll JF and De Groot AMB (Eds.), *Handbook of Bilingualism: Psycholinguistic Approaches*. Oxford: Oxford University Press, pp. 49–67.

Malt BC, Jobe RL, Li P, Pavlenko A and Ameel E (2016) What constrains simultaneous mastery of first and second language word use?. *International Journal of Bilingualism* **20**(6), 684–699. doi:10.1177/1367006915583565

Marecka M, Szewczyk J, Otwinowska A, Durlik J, Foryś-Nogala M, Kutyłowska K and Wodniecka Z (2020) False friends or real friends? False cognates show advantage in word form learning. *Cognition* **206**, 104477. doi:10.1016/j.cognition.2020.104477

Marian V, Blumenfeld HK and Kaushanskaya M (2007) The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language and Hearing Research* **50**, 940–967. doi:10.1044/1092-4388(2007/067)

Martin KI and Ellis NC (2012) The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition* **34**(3), 379–413. doi:10.1017/S0272263112000125

Martinez HDR (2015) *Analysing interactions of fitted models*. CRAN-R Project.

Matuschek H, Kliegl R, Vasishth S, Baayen H and Bates D (2017) Balancing Type I error and power in linear mixed models. *Journal of Memory and Language* **94**, 305–315. https://doi.org/10.1016/j.jml.2017.01.001

Michas IC and Henry LA (1994) The link between phonological memory and vocabulary acquisition. *British Journal of Developmental Psychology* **12**(2), 147–163. doi:10.1111/j.2044-835X.1994.tb00625.x

Mulík S, Carrasco-Ortiz H and Amengual M (2019) Phonological activation of first language (Spanish) and second language (English) when learning third language (Slovak) novel words. *International Journal of Bilingualism* **23**(5), 1024–1040. doi:10.1177/1367006918781061

Nation IS (2013) *Teaching & Learning Vocabulary*. Boston: Heinle Cengage Learning.

Odlin T (1989) *Language Transfer: Cross-Linguistic Influence in Language Learning*. Cambridge University Press.

Otwinowska A and Szewczyk JM (2019) The more similar the better? Factors in learning cognates, false cognates and non-cognate words, *International Journal of Bilingual* **22** (8), 974–991. doi:10.1080/13670050.2017.1325834

Otwinowska A, Foryś-Nogala M, Kobosko W and Szewczyk J (2020) Learning Orthographic Cognates and Non-Cognates in the Classroom: Does Awareness of Cross-Linguistic Similarity Matter? *Language Learning* **70**(3), 685–731. doi:10.1111/lang.12390

Peleg O, Degani T, Raziq M and Taha N (2020) Cross-lingual phonological effects in different-script bilingual visual-word recognition. *Second Language Research* **36**(4), 653–690. doi:10.1177/0267658319827052

Perfetti CA (2007) Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading* **11**, 357–383. doi:10.1080/10888430701530730

Prior A, Degani T, Awawdy S, Yassin R and Korem N (2017) Is susceptibility to cross-language interference domain specific?. *Cognition* **165**, 10–25. doi:10.1016/j.cognition.2017.04.006

R Core Team (2019) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, version 3.6.1. URL: https://www.R-project.org/.

R Core Team (2020) R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, version 4.0.3. Retrieved from www.R-project.org.

Raboyeau G, Marcotte K, Adrover-Roig D and Ansaldo AI (2010) Brain activation and lexical learning: the impact of learning phase and word type. *Neuroimage* **49**(3), 2850–2861. doi:10.1016/j.neuroimage.2009.10.007

Rice CA and Tokowicz N (2020) A review of laboratory studies of adult second language vocabulary training. *Studies in Second Language Acquisition* **42** (2), 439–470. doi:10.1017/S0272263119000500

Rodd JM, Berriman R, Landau M, Lee T, Ho C, Gaskell MG and Davis MH (2012) Learning new meanings for old words: effects of semantic relatedness. *Memory & Cognition* **40**(7), 1095–1108. doi:10.3758/s13421-012-0209-1

Saiegh-Haddad E and Henkin-Roitfarb R (2014) The structure of Arabic language and orthography. In Saiegh-Haddad E and Joshi R (eds), *Handbook of Arabic literacy*. Dordrecht, Netherlands: Springer, pp. 3–28. doi:10.1007/978-94-017-8545-7_1

Shatil E and Share DL (2003) Cognitive antecedents of early reading ability: A test of the modularity hypothesis. *Journal of Experimental Child Psychology* **86**(1), 1–31. doi:10.1016/S0022-0965(03)00106-1

Sonderegger M, Wagner M and Torreira F (2018) *Quantitative Methods for Linguistic Data*. v. 1.0. Retrieved from: http://people.linguistics.mcgill.ca/~morgan/book/

Sparks R, Patton J, Ganschow L, Humbach N and Javorsky J (2006) Native language predictors of foreign language proficiency and foreign language aptitude. *Annals of Dyslexia* **56**, 129–160. doi:10.1007/s11145-011-9335-6

Storkel HL, Maekawa J and Aschenbrenner AJ (2013) The effect of homonymy on learning correctly articulated versus misarticulated words. *Journal of Speech, Language, and Hearing Research* **56**(2), 694–707. doi:10.1044/1092-4388(2012/12-0122)

Tokowicz N and Degani T (2015) Learning second language vocabulary: Insights from laboratory studies. In Schwieter John W. (Ed.) *The Cambridge handbook of bilingual processing*. Cambridge, UK: Cambridge University Press, pp. 216–233.

Tomoschuk B, Ferreira VS and Gollan TH (2019) When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition* **22**(3), 516–536. doi:10.1017/S1366728918000421

Voeten C (2019) buildmer: Stepwise elimination and term reordering for mixed–effects regression. R package version 1.3. Retrieved from https://CRAN.R-project.org/package=buildmer

https://doi.org/10.1017/S1366728921000857 Published online by Cambridge University Press