LANGUAGE LEARNING
A Journal of Research in Language Studies

Language Learning ISSN 0023-8333

EMPIRICAL STUDY

Novel Word Learning Among Bilinguals Can Be Better Through the (Dominant) First Language Than Through the Second Language

Zoya Hirosh and Tamar Degani

University of Haifa

Abstract: When learning novel vocabulary in a third language (L3) through translations in the first language (L1), bilinguals may have more available cognitive resources and more accumulated experience in language regulation compared to when learning through translations in the second language (L2). In a study designed to test language of instruction (LOI) effects, 59 Hebrew–English bilinguals auditorily learned over two sessions 55 words in German, including three word types: cognates, overlapping in form and meaning between English and German; false cognates, overlapping in form but not meaning; and controls. Critically, half of the participants learned through their (dominant) L1 Hebrew, and half through their L2 English (which is also more similar to German). Results showed a significant LOI effect, with better learning through the (less similar) L1, especially for control items. Cognates were learned better in both LOIs, but false cognates were learned better relative to controls to a greater extent when the LOI was English. Together, results highlight the importance of LOI and item-based language similarity during multilingual novel word-learning.

Keywords vocabulary learning; language of instruction; multilingualism; language similarity

---

This work was funded by a Small Research Grant from the Faculty of Social Welfare and Health Sciences at the University of Haifa to Tamar Degani. The authors would like to thank Miri Goldberg and Margaryta Paliy for assistance with coding and analyses.

Correspondence concerning this article should be addressed to Tamar Degani, Department of Communication Sciences & Disorders, University of Haifa, Mount Carmel, Haifa, Israel. E-mail: tdegani@research.haifa.ac.il

The handling editor for this article was Theres Grüter.

DOI: 10.1111/lang.12457

Introduction

Novel language learning is a common task that extends beyond monolinguals learning a second language (L2) to bilinguals learning a third language (L3) or a fourth one. Several variables have been highlighted as affecting novel language learning, including the age of the learner (Birdsong, 2005), the context of learning (e.g., Linck, Kroll, & Sunderman, 2009), previous experience with language learning (e.g., Hirosh & Degani, 2018), the typological similarity of the novel language to the language(s) known (Rothman, 2011), the relationships among the words being learned, and the direction of translation during training and testing (Kemp & McDonald, 2021). However, little research has examined how the language of instruction (LOI) affects learning (though discussed by Bogulski, Bice, & Kroll, 2019 and by Tomoschuk, Duyck, Hartsuiker, Ferreira, & Gollan, 2021). For instance, when a Korean–English bilingual is learning Spanish, would learning Spanish through Korean, the first language (L1), be better or worse than learning through English (the L2)? The current project addresses this issue with respect to vocabulary learning and examines whether multilingual vocabulary learning is better when conducted through the L1 or through the L2.

Background Literature

Predictions for Language of Instruction Effects

Predicting an L1 Advantage

The effect of LOI on word learning may take one of several forms. First, learning of an L3 might be better through the L1 than through the L2. There are two main reasons to expect such an L1 advantage. For one, the *language regulation* proposal put forth by Bogulski et al. (2019) suggests that foreign language learning benefits from inhibition of the language from which learning takes place, and bilinguals have gained more experience in inhibiting the L1 in the process of learning and using the L2. Therefore, this proposal predicts better learning of an L3 through the L1 than through the L2. Support for this proposal comes from their study comparing bilingual and monolingual novel word learning, as discussed below.

Previous research had suggested that bilinguals outperform monolinguals in novel word learning (Kaushanskaya & Marian, 2009a, 2009b; for review, see Hirosh & Degani, 2018). For instance, Kaushanskaya and Marian (2009a) observed that English–Spanish and English–Mandarin bilinguals outperformed English monolinguals on a word-learning task. However, Bogulski et al. (2019) then suggested that this advantage is restricted to learning through the L1 but not through the L2. They first replicated the multilingual advantage in word learning, showing that English–Spanish bilinguals were more accurate than English monolinguals when learning Dutch words via English translations (the L1 for both groups). Critically, they then found that Spanish–English and Chinese–English bilinguals who learned the Dutch words via their *L2* (English) exhibited no advantage compared to monolinguals. Hence, Bogulski et al. (2019) concluded that the multilingual advantage is present only when the novel language is acquired via the native (or dominant) language. They reasoned that this has to do with bilinguals' experience in language regulation, namely in activation management of the language network via inhibitory control. Because inhibiting the L1 when speaking the L2 is required to a greater extent than inhibiting the L2 when speaking the L1 (Green, 1998; Guo, Liu, Misra, & Kroll, 2011), bilinguals become more skilled in regulating their L1 and inhibiting it than their L2. Hence, whereas English–Spanish bilinguals had extensive experience in inhibiting their L1 English during the use of their L2 Spanish, Spanish–English and Chinese–English bilinguals were less experienced in inhibiting their L2 English. As a result, when learning Dutch words through English, presumably requiring inhibition of English, the latter two groups of bilinguals did not demonstrate an advantage relative to monolinguals. A post hoc analysis further revealed that increased proficiency and dominance in the language through which learning took place (English) was associated with better learning, because English-dominant Spanish–English bilinguals outperformed Spanish-dominant Spanish–English bilinguals. Most relevant to the present study, these findings are in line with the suggestion that language learning depends on the LOI, and that performance differs between learning words through the L1 (or dominant language) and learning through the L2 (or less dominant language).

A second reason to expect better learning through the L1 is that speakers are typically more proficient in their L1 relative to their L2. Therefore, it might be the case that when proficiency in the language through which learning takes place is higher, there are more cognitive resources available for processing of the newly acquired information (Segalowitz, 2003; Tzelgov & Kadosh, 2009), especially early in learning. This *proficiency/resource-availability* proposal is supported by the general observed association between proficiency in the L1 and learners' ultimate proficiency in the L2 (e.g., Cummins, 1979), and more directly by a recent study in which proficiency in the language through which learning took place modulated learning (Degani & Goldberg, 2019). In that study, Degani and Goldberg compared native Hebrew speakers to Russian–Hebrew multilinguals in learning Arabic words and did not find a multilingual advantage when learning through Hebrew (the L2 of the multilinguals).

Instead, they found that learners' proficiency in the language through which learning took place (Hebrew) modulated learning, so that participants with higher Hebrew proficiency learned words in Arabic better than those with lower proficiency, regardless of their degree of multilingualism. Notably, in that study, Hebrew proficiency might have facilitated learning not only because this was the LOI and increased proficiency allowed for more cognitive resources, but also because Hebrew is more typologically similar to the novel language, Arabic. Thus, the results do not allow one to discern whether the facilitatory effect of Hebrew proficiency stemmed from its typological similarity to the foreign language or from its status as the LOI. As explained below, the current study addresses this issue directly by contrasting the contribution of language typological similarity with the contribution of language proficiency to the effect of LOI.

Importantly, because speakers' dominance often changes over time (Martin, Altarriba, & Kazanas, 2020), such that the L1 might not always be the more dominant (or proficient) language, it is not clear whether a learning advantage will emerge when the LOI is the first language in terms of the order of acquisition but not in terms of being the dominant language. Bogulski et al. (2019) suggested that for multilinguals who have experienced a switch in dominance such that they have become more dominant in their L2 rather than in their L1, a learning advantage may be observed when these multilinguals are learning through their dominant language, because they are experienced in regulating this dominant language in the course of using their first-acquired, less dominant language. However, in the study by Degani and Goldberg (2019) all participants learned the Arabic foreign language vocabulary through what was in fact their dominant language (Hebrew), even though it was the L2 for those Russian–Hebrew speakers, and yet no multilingual advantage emerged. It is therefore possible that the order of acquisition of the languages does play an important role and that a learning advantage for bilinguals over monolinguals is limited to cases where the LOI is the *first*-acquired language irrespective of its dominance (Degani & Goldberg, 2019). To prevent this complexity from potentially confounding the current study, as the first study to test the effects of LOI on learning within the same population, we ensured that participants' more dominant language was also their first-acquired one.

Predicting an L2 Advantage

Alternatively, learning might be better through the L2 because there is typically greater similarity between learning the L2 and L3 than between the L1 and L3 in terms of the age of onset, learning context and strategies, and degree of metalinguistic knowledge. Further, based on neurolinguistic accounts (e.g., Ullman, 2001), Bardel and Falk (2012) suggested that both the L2 and L3 as nonnative languages learned later than the L1, are more likely to rely more on declarative memory systems, whereas the L1 is more likely to rely more on procedural memory systems. If the L3 resembles the L2 more than the L1 in these regards, and one assumes that vocabulary learning can benefit from such resemblance, then learning through the L2 would be better than learning through the L1. Suggestive evidence in support of this prediction comes from studies conducted in the educational system, whereby a novel language is learned through the societal language by monolingual or bilingual children. For instance, Schwartz, Geva, Share, and Leikin (2007) examined the learning of English through Hebrew, which served as the L1 for some children but as the L2 for others (who had Russian as their L1) (see also, e.g., Schwartz, Kahn-Horwitz, & Share, 2014). A bilingual advantage was observed, attributed to enhanced experience of the Russian-Hebrew bilinguals in learning to read and to the typological similarity of their L1 Russian with the to-be-learned L3 English. Interestingly, the fact that the bilinguals outperformed monolinguals (those with L1 Hebrew) may also imply that learning through Hebrew when it is the L2 (for bilingual children) is better than learning through Hebrew when it is the L1 (for monolingual children). These studies therefore raise the possibility that learning via the L2 may yield a learning advantage, but no strong conclusions can be drawn due to the confounding role of multilingualism, language similarity, and cross-population comparisons.

Predicting no L1/L2 Effect

A third possible predictive pattern is that language order per se does not affect learning, but instead it is language typological similarity alone that influences performance (e.g., Schwartz et al., 2007, 2014). The role of language typological similarity has been emphasized in the literature on cross-language influences in learning (for review, see, e.g., Ringbom & Jarvis, 2009). For instance, Schepens, van der Slik, and van Hout (2016) found that the distance between learners' prior languages (L1 and L2) and L3 (Dutch) explained almost half of the variance in performance on a L3 speaking proficiency test among a sample of over 40,000 multilinguals with various L1s and L2s. Given the important role of language similarity, one may predict that learning would be facilitated through a more typologically similar language, whether it be a L1 or a L2. The present study will therefore examine the role of language similarity, manipulating LOI, to reveal potential interactions between the two factors in the case of vocabulary learning. Further, the current study goes beyond whole-language typological similarity to examine how the similarity of specific items interacts with the effects of LOI and typological similarity.

Thus far, no studies that we know of have compared learning through the L1 versus the L2 within the same population, with the exception of a recent study by Tomoschuk and colleagues (Tomoschuk et al., 2021, Experiment 2). In that study, learning was not the relevant outcome measure, because the focus was more on the engagement of inhibition after learning. Dutch–English bilinguals learned novel L3 words in an artificial language (called "Ibararpa") that was based on Italian pseudowords, either via their L1 Dutch or via their L2 English. Participants who learned via the L1 Dutch scored slightly higher (though not significantly so) on the final retrieval block compared to those who learned via the L2 English. In another experiment using the same learning paradigm (Tomoschuk, 2019, Experiment 3), a within-participant design was used so that Spanish–English bilinguals learned novel Hebrew words, with half the words learned via their L1 Spanish and half via their L2 English. The difference between words learned through the L1 versus the L2 was again nonsignificant, although in this case learning through the L2 English was associated with slightly higher performance than learning through the L1 Spanish.

Notably, in both experiments the L1 and the L2 were either equally similar to the L3 (Dutch–English vs. Ibararpa) or equally typologically different from the L3 (Spanish–English vs. Hebrew), excluding the role of language similarity. In the present study, the to-be-learned novel language resembled one of the languages of the speakers but not the other. Thus, our study contrasted the contribution of language similarity with that of LOI.

Item-Based Versus Whole-Language Similarity

In the present study, Hebrew (L1)–English (L2) bilinguals learned German as a L3. We chose German because of its typological relatedness to speakers' L2 English, both being Germanic Indo-European languages (Hawkins, 1986; see also Hammarström, Forkel, Haspelmath, & Bank, 2020), and its dissimilarity to Hebrew (a Semitic Afro-Asiatic language). Critically, in addition to the similarity between English and German at the typological whole-language level, similarity was further manipulated at the item level. Participants learned three types of German words differing in the degree of lexical-form and semantic overlap with English: (a) cognate words, which overlap in both lexical form and meaning between the two languages (e.g., *Maus* in German translates as "mouse" in English); (b) false cognates (FCs), which overlap in lexical form but not in meaning (e.g., *Gift* in German translates as "poison" in English but is pronounced like "gift"); and (3) control words (noncognates), which do not overlap in lexical form across languages (e.g., *Saft* in German translates as "juice" in English). Previous studies have shown that cognate words are easier to learn than control words (De Groot & Keijzer, 2000; Elias & Degani, under review; Lotto & de Groot, 1998; and see evidence in Otwinowska, Foryś-Nogala, Kobosko, & Szewczyk, 2020), and thus we expected a cognate facilitation effect.

In contrast, although FCs are typically more difficult to process than control words for proficient bilinguals (see Degani & Tokowicz, 2010b, for a review), studies that directly tested the patterns of acquiring FCs have shown contradicting evidence. Specifically, due to the similarity in lexical form across languages, facilitation may be observed compared to control items (Mulík, Carrasco-Ortiz, & Amengual, 2019). Evidence for this suggestion comes from within-language vocabulary learning, in which learning of a new meaning for a known word is compared to learning of novel words (e.g., Fang & Perfetti, 2017). Such learning of a new meaning for a known word resembles FC learning in the need to map a novel meaning to the same known form. Fang and Perfetti (2017) observed an advantage for learning new meanings for known forms over learning of novel words, and this may imply that FCs should be learned better than control words.

On the other hand, learning a new meaning for a known form (e.g., that *Gift* means "poison") may exert meaning competition and cause difficulty in learning FC items (Rodd et al., 2012). Indeed, in a recent study FC words were found to be less learnable than control (and cognate) words (Otwinowska & Szewczyk, 2019). However, other research observed no difference between FC and control words (Elias & Degani, under review; Ghazi-Saidi & Ansaldo, 2017; Otwinowska et al., 2020). For instance, Elias and Degani (under review) had adult Hebrew speakers learn cognate, FC, and control Arabic words as novel words. Results showed comparable overall performance on FC and control words. The authors interpreted this null effect as reflecting the operation of two opposing mechanisms of form facilitation and meaning competition that are both at play and may cancel each other out. Presumably, when overall performance was examined the inhibitory effect of meaning competition counteracted the facilitative effect of form overlap, resulting in no overall effect. To support their proposal that both processes are at play, Elias and Degani further examined how individual differences modulated the difference in learning of FC and control words. They found that the effect was modulated by participants' characteristics, such that participants with higher phonological short-term memory and those with lower Hebrew verbal fluency did show an advantage for FC over control words. Elias and Degani suggested that participants with higher phonological short-term memory rely more on the phonological form and thus benefit more from the form overlap of FCs. At the same time, participants with lower L1 verbal abilities rely less on meaning than individuals with higher L1 abilities and are thus less affected by the competition in meaning from the L1. The present study examines the possible existence of these two opposing mechanisms in the learning of FCs using a different to-be-learned language (German) when form overlap is manipulated with learners' L2 (English) and not their L1 (Hebrew).

Because in the current study form overlap is manipulated with one of the languages of the bilingual learners (English but not Hebrew), we expect item-based similarity to exert stronger influences when participants learn directly through that language (English). In such a condition, the form similarity is more transparent, and learners may attend to it more, utilizing their prior knowledge to achieve better learning. Notably, however, it is possible that participants would benefit from the similarity between English and German even when they are learning through Hebrew. Indeed, Bartolotti and Marian (2017) had English–German bilinguals learn orthographically presented novel words that were similar to English words, to German words, or to both. Performance was comparable for novel words that overlapped with one of the bilinguals' languages (English or German) and those that overlapped with both, suggesting flexible transfer of L1 and L2 knowledge to the novel words (Bartolotti & Marian, 2017). Note, however, that whereas in that study novel words were learned in association with pictures, allowing bilinguals to use both of their existing languages during learning, in the current study novel words are learned in association with translations in one language only. Thus, transfer is predicted to be greater for participants learning through English, because of higher activation of a language similar to the novel language, which may allow learners to attend to similar characteristics of the novel words.

The Current Study

The current study examined word learning among multilingual speakers, asking whether vocabulary learning in a L3 was better through the L1 or through an L2 more similar to the L3. Hebrew–English bilinguals learned novel words either through their L1 (Hebrew) or their L2 (English). Participants completed a two-session learning and testing procedure, learning 55 novel words in German that differed from English in their degree of similarity in form and meaning so as to include cognates, FCs, and control words. This word-type manipulation allowed examination of language similarity at both the global (whole-language) level and the item level.

With respect to LOI, we predicted better learning through participants' L1 (Hebrew), under the assumption that bilinguals are more experienced at inhibiting their L1 than they are experienced at inhibiting their L2 and that inhibition of the LOI promotes learning. Further, we also predicted better learning through the L1 under the assumption that the enhanced proficiency in the L1 would free up more cognitive resources required for the learning process. These LOI effects were thus expected to be driven by participants' proficiency profile (L1 vs. L2) and not by the whole-language typological similarity of the languages, which would have predicted better learning through their L2 (English) more similar to German than through their L1 (Hebrew) less similar to German. Notably, learning of control words, which did not overlap in terms of form and meaning across languages, provided the cleanest test of this hypothesis. If learning of control words proved to be better through Hebrew compared to through English, although the latter shares more commonalities with German, the results would support the role of proficiency profile over typological similarity in driving LOI effects.

Although we predicted that whole-language typological similarity would not determine L3 learning, we expected item-based similarity to affect learning. In particular, we predicted that, at the item level, cognates would be easier to learn than control words and that this cognate facilitation effect would be greater for the group learning through the L2 because the similarity at the whole-language level in this condition could highlight the similarity at the item level. Conversely, as we discussed above, the difference between FCs and control words might reveal a FC advantage or a FC disadvantage, and we hypothesized this difference to be larger for the group learning through the more similar L2.

Method

Participants

Sixty-six Hebrew–English bilingual adults participated in the study. They were native Hebrew speakers who had learned their L2 English in school, with no significant immersion experience in an English-speaking environment, no exposure to any other language at home, and no prior knowledge of German (or Dutch). Within this population, English instruction typically begins in the third grade, starting with an emphasis on oral language proficiency, and progresses to also include a focus on the written form. At the time of testing, participants were university students who had reached at least moderate proficiency in English as a prerequisite for academic studies. Participants' L1 and L2 proficiency were assessed via objective and subjective language measures (see

Table 1). They reported no attention disorders or learning disabilities. They were compensated with class credit or payment for participation.

Participants were randomly divided into two learning conditions, one group learning German words through their L1 (Hebrew) and the other through their L2 (English). Additional measures, including phonological short-term memory, inhibitory control, working memory, Hebrew and English proficiencies, and a vocabulary posttest, were collected to verify that there were no baseline differences between the two experimental groups. Indeed, the two groups were well matched on all measures with the exception of age, which was covaried in the analysis, and the self-report percentage of reading in Hebrew and English and talking in English (see Table 1). Because the language proficiency and use questionnaire was administered following experimental tasks that differed in language across groups, we suspect that these differences were the result of an influence of the experimental task on the self-report measure.

Data from seven participants were excluded, two because of exposure to an additional language at home, three because they were born or had lived in an English-speaking country, one because of very low English proficiency (as assessed by the vocabulary posttest), and one because of failure to return for the second experimental session. Analyses were therefore based on a final set of 59 participants.

Sample size was determined based on previous vocabulary learning studies in which item type and participants' characteristics were observed (Degani & Goldberg, 2019; 30 participants per group, 12 items per condition). Nonetheless, we acknowledge that the estimated power to detect an interaction (of Cohen's $d = 0.3$) between LOI and item type with 29 participants per group and 15 items per condition is rather low (power = 0.531, using PANGEA; Westfall, 2016).

Materials

All learning and testing materials (Hirosh & Degani, 2021a, 2021b, 2021c, 2021e) are available on IRIS (iris-database.org) and the OSF platform (https://osf.io/habuz/?view_only=bde3ec16027c4671a13e264913ee3d99). They include language history questionnaire, recorded stimuli, and vocabulary posttest, as well as the experimental scripts. Fifty-five critical German words served as stimuli in the learning task (see Table 2 for example stimuli and Appendix S1 in the Supporting Information online for all items). German words were recorded by a native German speaker and were presented auditorily along with their orthographic Hebrew or English translations. These included 15 cognate words, overlapping in both lexical form and meaning with English; 15 FC words,

Table 1 Means (standard deviations) and independent samples *t* tests for linguistic and cognitive characteristics by language of instruction group

Measure	L1 Hebrew ($n = 30$)	L2 English ($n = 29$)	$t(57)$	p
Age (years)	24.50 (3.25)	26.69 (4.74)	−2.08	**.042**
Education (in years)	13.97 (1.63)	14.38 (1.78)	−0.93	.357
Socioeconomic status (maternal edu. in years)	15.23 (3.66)	14.62 (2.73)	0.73	.470
Age began learning English	7.25 (1.74)	7.59 (2.10)	−0.66	.513
Rating of Hebrew proficiency (0–10 scale)[a]	9.60 (0.56)	9.66 (0.67)	−0.34	.733
Rating of English proficiency (0–10 scale)[a]	7.57 (1.00)	7.83 (1.20)	−0.92	.361
Rating of Hebrew use (0–10 scale)[b]	7.78 (1.09)	7.88 (1.67)	−0.30	.766
Rating of English use (0–10 scale)[b]	6.40 (1.43)	6.49 (1.75)	−0.22	.831
Reading in Hebrew (%)[c]	82.80 (15.01)	72.62 (19.46)	2.26	**.028**
Reading in English (%)[c]	16.93 (12.77)	26.93 (18.99)	−2.35	**.023**
Talking in Hebrew (%)[c]	87.37 (16.80)	82.21 (17.09)	1.17	.247
Talking in English (%)[c]	11.39 (12.96)	20.81 (18.90)	−2.16	**.037**
Exposure to Hebrew (%)[c]	76.63 (11.69)	73.24 (11.10)	1.14	.258
Exposure to English (%)[c]	19.60 (8.83)	24.97 (11.66)	−1.99	.052
Number of languages	2.77 (1.07)	2.45 (0.57)	1.43	.160
Rating of L3 proficiency (0–10 scale)[a]	3.55(2.02)	2.92 (1.93)	0.76	.453
Rating of L3 use (0–10 scale)[b]	1.70 (1.73)	1.44 (1.27)	0.41	.685
Hebrew semantic fluency (mean per minute)	21.97 (5.82)	21.41 (5.25)	0.38	.703
English semantic fluency (mean per minute)	14.47 (5.15)	14.90 (4.67)	−0.34	.738
Phonological memory (possible range 0–14)	5.83 (1.68)	5.41 (1.15)	1.11	.270
Working memory (possible range 0–21)	11.67 (2.95)	11.10 (2.02)	0.85	.395
Inhibitory control (bin score)	404.80 (72.70)	422.93 (89.33)	−0.86	.395
Unknown English items[d]	1.27 (1.60)	1.24 (1.50)	0.06	.950

Note. Values of p in boldface represent those for comparisons where $p < .05$ for group differences.

[a]Language proficiency in reading, writing, conversation, and speech comprehension (0 = lowest ability; 10 = highest ability).

[b]Language use in reading, writing, conversation, internet, listening, and TV watching (0 = least use; 10 = highest use).

(*Continued*)

Table 1 (Continued)

[c] Ratings for reading, talking, and exposure in the known languages sum to 100% for each participant for each set of ratings.
[d] The means of unknown English items are based on the vocabulary posttest including the 55 experimental items.

Table 2 Examples of stimuli for each word type and language of instruction

Word type	German	English	Englishfalse translation	Hebrew
Cognate	Maus /maʊs/	Mouse /maʊs/		עכבר /ʕaxbar/
False cognate	Gift /gɪft/	Poison /pɔɪzən/	Gift /gɪft/	רעל /raʕal/
Control item	Saft /zaft/	Juice / ʤus /		מיץ /mits/

overlapping with English (but not Hebrew) in lexical form but not in meaning; and 25 control words that did not overlap with English (or Hebrew) in lexical form.

Stimuli Selection Procedures

Cognate and FC words were selected to overlap phonologically between English (L2) and German, but not between Hebrew (L1) and German. Phonological similarity was informed by (a) a subjective norming study with a group of 20 Hebrew speakers who did not participate in the main study (see Appendix S2 in the Supporting Information online for details) and (b) an objective computation of Levenshtein's phonological distance based on a detailed phonological representation using the PanPhon 0.7 Python package (https://pypi.org/project/panphon/0.7/), which applies a set of rules for adding diacritics and modifiers to International Phonetic Alphabet segments based upon their phonological features. Comparisons across word types revealed that cognates and false English translations of the FCs scored significantly higher in the German–English objective and subjective similarity measures than did the actual translations of the FCs and controls but did not differ from each other. The phonological overlap across English and German was not complete for these items, but, critically, all cognates and false translations of FCs were judged as at least a 3.5 on a 1 to 5 scale of phonological similarity, whereas all control items and correctly translated FCs were judged as less than a 2 on the same scale (see Appendix S2 in the Supporting Information online for details).

Further, there were no differences across the three types of English word in terms of their length (in number of letters and syllables in German, English, and Hebrew), age of acquisition (Brysbaert & Biemiller, 2017), concreteness, and imageability (the latter two both based on the MRC Psycholinguistic Database; Wilson, 1988) (all ps > .05; see Appendix S3 in the Supporting Information online). However, a significant difference in frequency in English and German (based on the SubtLex corpus via Clearpond), as well as in Hebrew (based on the heTenTen 2014 corpus via SketchEngine; see Kilgarriff et al., 2014), was found, such that cognates were more frequent than FCs and controls in all three languages (ps < .05), with no difference between FCs and controls (ps > .05). Frequency was therefore included as a covariate in the analyses.

Procedure

Each participant learned all 55 German words (15 cognates, 15 FCs, and 25 controls), either through the L1 Hebrew or through the L2 English, manipulated between participants. All instructions and communication with the experimenter (a Hebrew-English bilingual) matched the LOI, such that participants who learned through L1 Hebrew received the instructions in Hebrew and communicated with the experimenter in Hebrew, whereas those who learned through the L2 English, received English instructions and interacted with the experimenter in English. Learning and testing took place over two sessions separated by 2 days (see Table 3). All computerized tasks were administered via E-Prime 2 (2019).

Learning Trial Procedure

The learning phase in the first session consisted of two types of training cycles (following Degani & Goldberg, 2019): a L1/L2-to-L3 word repetition procedure and a L3 retrieval (recall) attempt. In the first type of learning cycle, a fixation cross appeared at the center of a screen for 750 milliseconds followed by a 250-millisecond blank screen, which was then replaced by either the Hebrew (L1) or the English (L2) written translation (depending on the LOI condition), presented for 1,000 milliseconds, and followed by the auditory presentation of the German (L3) word via headphones. A question mark then appeared, and participants were instructed to repeat the German word aloud.

Table 3 Overall procedure for the German word learning and testing

Tasks in Session 1	Tasks in Session 2 (2 days later)
Hebrew proficiency	Test: translation recognition
Learning cycle 1: Hebrew/English to German repetition	Learning cycle 2: Hebrew/English to German retrieval attempt
Learning cycle 2: Hebrew/English to German retrieval attempt	Test: Hebrew/English to German translation production
Test: translation recognition	Inhibitory control
English proficiency	Phonological short-term memory
Language history questionnaire	Working memory
	Vocabulary posttest

An interstimulus interval of 1,000 milliseconds of blank screen was used. Each of the 55 novel words was presented once in a random order.

In the second type of training cycle, after seeing the L1 or L2 translation, a screen with a question mark appeared and participants were asked to try to produce the German word. Following this retrieval attempt, participants heard the correct form of the German word. Such retrieval attempts have been shown to strengthen learning (Kang, Gollan, & Pashler, 2013; see Tokowicz & Degani, 2015, for a review). In this type of training cycle, each word was presented twice in a random order for a total of 110 training trials.

In the second session, 2 days later, participants completed only the second (retrieval) type of training cycle.

Testing Trial Procedure

Participants were tested via a translation recognition test and a forward translation production test (Degani & Tokowicz, 2010a; Degani, Tseng, & Tokowicz, 2014). In both tests, trial order was randomized for each participant by the computer program. In the translation recognition test, participants heard a German word followed by a visual presentation of four possible Hebrew or English translations, according to the LOI condition, and were asked to select the correct translation of the German word by pressing one of four buttons (marked with 1 through 4) on the response box with their dominant hand as quickly and accurately as possible. The nontarget alternative translations were sampled from the set of 55 translations used in the study. In the forward translation production test (L1/L2-to-L3), participants saw a Hebrew or English word, according to their LOI condition, and were asked to provide its German translation out loud as quickly and accurately as possible. Reaction times were measured between the onset of the prompt and the onset of the vocal response, using a voice-key trigger, and responses were recorded for later coding of accuracy. Partially correct pronunciations in which vowels or consonants were omitted or added (e.g., /Glock/ instead of /Glocke/) were coded as incorrect trials. Slight changes in accent for nonnative phonemes were considered correct. A single researcher coded the responses, and a second coder coded 10% of the data. Cohen's kappa for interrater reliability computed on this basis indicated substantial agreement ($\kappa = .79$).

Controlling for Cognitive and Linguistic Differences Across Language of Instruction Groups

Phonological short-term memory task. A version of the nonword repetition task (e.g., Yoo & Kaushanskaya, 2012) was used in Hebrew (Shatil & Share, 2003) to estimate participants' phonological short-term memory. Participants were asked to repeat out loud in the same order presented sets of Hebrew nonwords ranging in set sizes of two to eight words. The test ended when the participant failed to accurately repeat both sets of a given length. Although split-half reliability with Spearman–Brown correction for this task has been reported as .67 in previous research (Degani & Goldberg, 2019), in the current sample, it was only .48.

Working memory task. The Number–Letter sequencing task (a subtest of the Wechsler Adult Intelligence Scale Version III – WAIS-III; Wechsler, 1997), in its Hebrew version (WAIS-III HEB; http://www.psychtech.co.il), was used to estimate participants' verbal working memory span. Participants were presented with sets of varying sizes of two to eight sequences of letters and numbers. They were asked to first repeat the numbers in ascending order and then repeat the letters in their alphabetic order (e.g., a seven-character string such as 3esa41b should be repeated back as 134abes). The test ended when participants failed to accurately repeat any of the three strings of a given length. Split-half reliability with Spearman–Brown correction for this task was .67.

Inhibitory control. A Numeric Stroop task (Prior, Degani, Awawdy, Yassin, & Korem, 2017, adopted from Hernández, Costa, Fuentes, Vivas, & Sebástian-Gallés, 2010) was used to measure participants' inhibitory control. Participants indicated by button press, as quickly and accurately as possible, how many items appeared on the screen (range from 1 to 3). There were three experimental conditions, with trials presented in a random order. In the congruent condition, digits were used as items, and their numeric value was congruent with the number of items (e.g., 333 or 22). In the incongruent condition, the numeric value of the digits was not the same as the number of digits (e.g., 33 or 222). In the control condition, nondigits were used as items, creating no conflict between the number of items and the numeric value of the items (e.g., xxx). Inhibition was calculated as the difference in performance between congruent and incongruent trials, using the bin scoring method (as described by Hughes, Linck, Bowles, Koeth, & Bunting, 2014), which has been shown to improve reliability (Cronbach's alpha = .78 in Prior et al., 2017). Cronbach's alpha in the current study was .78. Higher bin scores represent larger costs.

Hebrew and English proficiency. A semantic category fluency task (as used by Kavé, 2005) was used to measure participants' oral proficiency in Hebrew and English. Participants were asked to produce as many different Hebrew or English words that belong to a certain semantic category as possible within a 1-minute time limit. Two different semantic categories for each language were counterbalanced across participants: one wide category (animals or fruits and vegetables) and one narrow category (occupations or furniture). Because preliminary analysis revealed significant differences between the two narrow categories (furniture being smaller than occupations), a fluency score was extracted based on only the wider category in each language. Reliability measures were thus not available for this task, as there was only one score for each participant (i.e., the number of items produced in the wider category).

Language history questionnaire. Participants completed a modified version of the LEAP-Q (Marian, Blumenfeld, & Kaushanskaya, 2007), in which they provided information on their proficiency, use, and learning circumstances for the languages that they knew.

Vocabulary posttest. To ensure participants were familiar with the English translations of the learned German words, participants received a list of Hebrew–English translations (corresponding to the novel German words) and were asked to indicate which English words or translation pairs they did not know. Unknown items were excluded from analysis for each participant (less than 3% of the data), and the number of exclusions did not differ by group (see Table 1).

Results

All relevant data (Hirosh & Degani, 2021d), as well as analysis scripts (Hirosh & Degani, 2021f) are available on both IRIS (iris-database.org) and the OSF platform (https://osf.io/habuz/?view_only=bde3ec16027c4671a13e264913ee3d99). Table 4 presents the observed means with standard errors for the percentage of errors and for reaction times (RTs) on correct responses in each test (translation recognition in the first session, translation recognition in the second session, translation production in the second session)

Table 4 Observed means (standard errors) for the percentage of errors (top) and for reaction times on correct responses (bottom), as a function of task, language of instruction, and word type

	Translation recognitionSession 1		Translation recognitionSession 2		Translation productionSession 2	
Word type	Hebrew	English	Hebrew	English	Hebrew	English
			Percentage of errors			
Cognate	0.03 (0.01)	0.02 (0.01)	0.01 (0.01)	0.02 (0.01)	0.09 (0.02)	0.11 (0.02)
FC	0.16 (0.02)	0.23 (0.02)	0.18 (0.02)	0.25 (0.02)	0.48 (0.03)	0.63 (0.03)
Control	0.21 (0.01)	0.33 (0.02)	0.23 (0.02)	0.36 (0.02)	0.59 (0.02)	0.77 (0.02)
			Reaction times (milliseconds)			
Cognate	1,365 (36)	1,285 (30)	1,344 (37)	1,338 (31)	1,160 (30)	1,067 (24)
FC	2,285 (88)	2,294 (77)	2,389 (82)	2,685 (89)	1,758 (60)	1,644 (83)
Control	2,595 (71)	2,659 (68)	2,782 (77)	2,783 (73)	1,987 (68)	1,728 (64)

Note. Standard errors calculated over all data points including within-participant variables following Morey (2008), using the function described by Chang, W. (n.d.) http://www.cookbook-r.com/Graphs/Plotting_means_and_error_bars_(ggplot2). FC = false cognate.

for the final sample of 59 participants as a function of LOI (Hebrew vs. English) and word type (cognate, FC, control). RTs on correct responses were trimmed to remove trials on which latencies were more than 2.5 standard deviations from the mean of each participant on correct responses (excluding about 4% of the production data and about 3% from the recognition data).

Data Analysis Approach

Overall error rates and RTs were analyzed using linear mixed-effects models because these models allow one to simultaneously account for variance related to participants and to items. Error rate data were analyzed following a binomial distribution (i.e., mixed effects logistic regression). RTs were log-transformed prior to analyses to reduce skew in the distribution. The models included dummy-coded fixed effects of LOI as a between-participant variable (Hebrew vs. English, with English set as the reference), word type as a within-participant variable (cognate, FC, control, with control set as the reference), and the interaction between LOI and word type. For the recognition test, session (Session 1 vs. Session 2, with Session 1 set as the reference) was also included as a within-participant variable and allowed to interact with LOI and word type. To control for differences across the LOI groups, participants' normalized age was included as a covariate. Similarly, to control for differences across word types in the frequency of the Hebrew or English translations, and because these were based on different corpora for English and Hebrew stimuli, we computed a standardized score within each language based on log frequency. These standardized translation frequency scores were included as covariates in the maximal model.

The random structure of the model included by-participant and by-item intercepts, as well as by-participant slope for word type, and by-item slope for LOI (and session in the recognition test). These maximal models were fitted using the buildmer function in the buildmer package (Version 1.3; Voeten, 2019) in R (Version 3.6.1, R Core Team, 2019), which uses the (g)lmer function from the lme4 package (Version 1.1.-21; Bates, Mächler, Bolker, & Walker, 2015). Using backwards stepwise elimination, the buildmer function starts from the most complex model and systematically simplifies the random structure until the model converges. Once the maximally converging model has been identified, the function estimates the significance of all fixed effects based on Satterthwaite degrees of freedom using the lmerTest package (Version 3.1-0; Kuznetsova, Brockhoff, & Christensen, 2017) for the RT analyses, or the Wald degrees of freedom for the error rate following the 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 (Version 0.2-1; De Rosario-Martinez, 2015) for the χ^2, the contrast function from the emmeans package (Version 1. 5. 2-1; Lenth, 2020) for the SE, and the tab_model function from the sjPlot package (Version 2.8.7; Lüdecke, 2021) with Bonferroni adjustments for multiple comparisons. Selected models are presented in Appendix S4 in the Supporting Information online. Model summaries (obtained from the summary function) are presented in Tables 5, 7, 9, and 11, and in Appendix S5 in the Supporting Information online. Note that, because fixed effects were dummy-coded, the effects presented in these tables reflect simple effects rather than main effects (e.g., the coefficients for LOI reflect its effect at the reference level of the other variables, and not across all levels). The main effects of each fixed variable were obtained from the anova function and are presented in the text. η_p^2 was obtained from the anova_stats function from the sjstats package (Version 0.18.1, Lüdecke, 2021). Significance was evaluated with an alpha level of 0.05, with Bonferroni adjustments for multiple comparisons when necessary. R^2 were used as a measure of effect size of the models, computed using the MuMIn package (v. 1.43.17, Bartoń, 2020).

Translation Recognition

The error rate analysis revealed that performance did not significantly change from Session 1 to Session 2, so this variable was not maintained in the selected model (see Table 5). Critically, there was a significant effect of LOI, $F(1) = 6.5$, $p = .001$, such that learning through Hebrew was less error-prone than learning through English. In addition, there was a main effect of word type, $F(2) = 147.15$, $p < .001$. Paired comparisons with Bonferroni corrections revealed that cognates elicited significantly fewer errors than FCs and controls, with marginally fewer errors for the FCs compared to the control items (see Table 6). Although the interaction between LOI and word type was not significant, the patterns of means suggested differential effects (for data-rich figures, see Figure 1 and Appendix S6 in the Supporting Information online). Due to our theoretical interest, we examined performance in each LOI separately and each word type separately (see Appendix S5 in the Supporting Information online). These analyses revealed that, whereas in the English (L2) instruction condition control items elicited more errors than FCs, in the Hebrew (L1) instruction condition this difference did not reach significance (see Table 6). In addition, there was no effect of LOI for cognates or FC items, but a significant advantage was observed for learning through Hebrew relative to learning

Table 5 Model summary predicting error rate in the translation recognition test as a function of session, language of instruction (LOI), and word type

Fixed effects	*b*	*SE*	95% CI	*z*	*p*
Intercept	−0.86	0.23	[−1.30, −0.41]	−3.76	<.001
Session (2nd)	–	–	–	–	–
LOI (Hebrew)	−0.67	0.27	[−1.20, −0.15]	−2.54	.011
Word type (cognate)	−3.37	0.28	[−3.92, −2.83]	−12.18	<.001
Word type (FC)	−0.51	0.22	[−0.93, −0.09]	−2.36	.018

Random effects	Variance	*SD*	*r*
Participant (intercept)	0.89	0.95	
Word type (cognate)	–	–	–
Word type (FC)	–	–	–
Session (2nd)	–	–	–
Item (intercept)	0.45	0.67	
LOI (Hebrew)	0.12	0.34	−.56
Session (2nd)	–	–	–
Residual			
$R^2_{\text{marginal}} = .33$; $R^2_{\text{conditionnal}} = .51$			

Note. Fixed effects reflect simple effects relative to the reference level when other variables are at their reference level without correction for multiple comparisons. For main effects, see *F* values in the text. FC = false cognate; – is used when data were not obtained because the effect was not retained in the selected model.

Table 6 Summary of odds ratios of error rates for pairwise comparisons with Bonferroni corrections for multiple comparisons for the word type effect in the translation recognition test within each language of instruction (LOI) group separately

Group	Comparison	*OR*	*SE*	95% CI	*z* ratio	*p*
Across LOI	Cognates vs. controls	29.15	8.08	[15.02, 56.58]	12.18	<.001
	FCs vs. controls	1.66	0.36	[0.99, 2.78]	2.36	.055
	Cognates vs. FCs	0.06	0.02	[0.03, 0.12]	−9.75	<.001
Hebrew	Cognates vs. controls	16.48	8.32	[4.92, 55.16]	5.55	<.001
	FCs vs. controls	1.29	0.33	[0.71, 2.37]	1.02	.921
	Cognates vs. FCs	0.08	0.04	[0.02, 0.25]	−5.26	<.001
English	Cognates vs. controls	42.35	14.83	[18.32, 97.93]	10.70	<.001
	FCs vs. controls	1.89	0.47	[1.05, 3.42]	2.60	.028
	Cognates vs. FCs	0.05	0.02	[0.02, 0.11]	−8.35	<.001

Note. FCs = false cognates. Three multiple comparisons were used to calculate the Bonferroni adjustment.

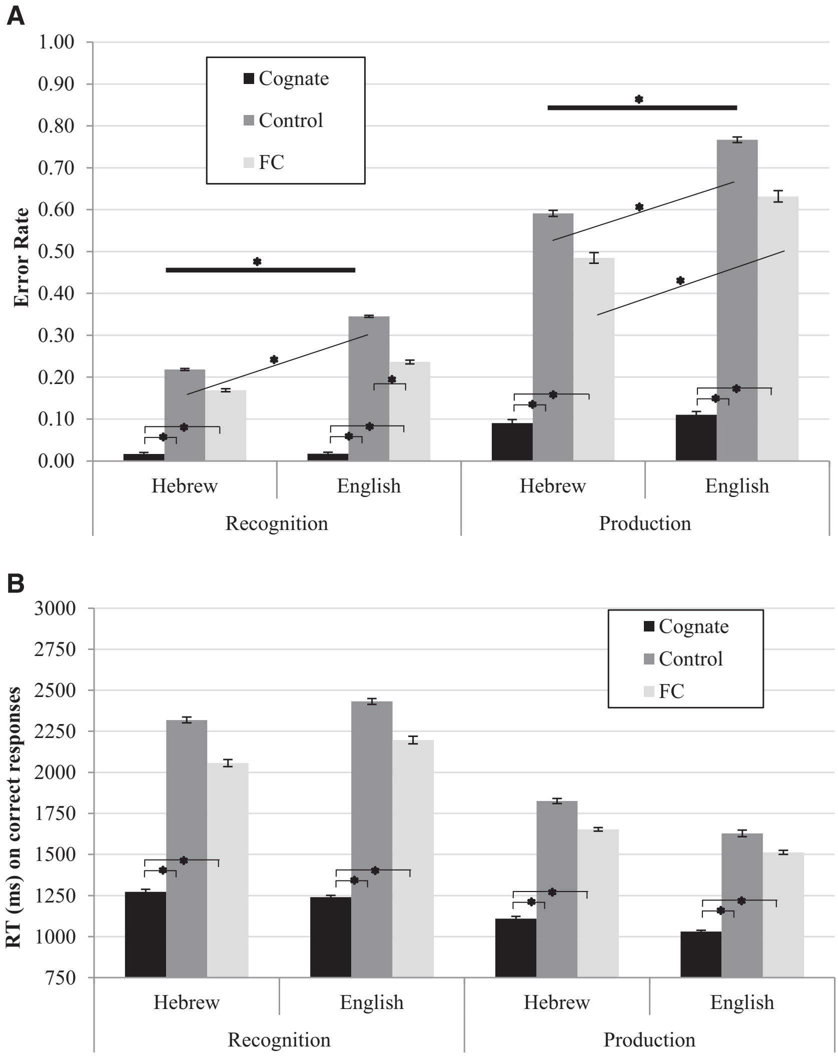

Figure 1 Estimated error rates (a) and reaction times (RTs) on correct responses (b) as a function of task, word type and language of instruction group (error bars represent standard errors calculated for within-participant variables following Morey, 2008). FC = false cognate. Predicted means are based on a model including the interaction between word type and language of instruction, excluding session in the recognition task. Means are back-transformed to raw RTs.

Table 7 Model summary predicting log reaction time in the translation recognition test as a function of session, language of instruction (LOI), and word type

Fixed effects	*b*	*SE*	95% CI	*df*	*t*	*p*
Intercept	3.35	0.03	[3.30, 3.41]	120.01	123.10	<.001
Session (2nd)	0.02	0.01	[0.01, 0.04]	53.24	2.82	<.001
LOI (Hebrew)	0.01	0.02	[−0.04, 0.06]	64.64	0.32	.75
Word type (cognate)	−0.27	0.03	[−0.33, −0.22]	71.92	−10.19	<.001
Word type (FC)	−0.04	0.03	[−0.09, 0.01]	63.56	−1.57	.122

Random effects	Variance	*SD*	*r*	*r*	*r*
Participant (intercept)	0.02	0.15			
Word type (cognate)	0.01	0.09	−.88		
Word type (FC)	0.00	0.06	−.44	.54	
Session (2nd)	0.00	0.03	−.11	−.04	.32
Item (intercept)	0.01	0.08			
LOI (Hebrew)	0.00	0.04	−.25		
Session (2nd)	0.00	0.03	−.09	−.43	
Residual	0.04	0.19			
$R^2_{\text{marginal}} = .21$; $R^2_{\text{conditionnal}} = .49$					

Note. Fixed effects reflect simple effects relative to the reference level when other variables are at their reference level without correction for multiple comparisons. For main effects, see *F* values in the text. FC = false cognate. *rs* reflect the correlation matrix among the random effects.

through English for control items, $F = 8.19, p = .004$ (see Appendix S5 in the Supporting Information online).

In the RT analyses, there was a significant effect of session, such that RTs increased from Session 1 to Session 2, $F(1, 53.24) = 7.96, p < .001, \eta_p^2 = .002$ (see Table 7). Further, there was a significant effect of word type, $F(2, 66.55) = 55.52$, $p < .001$, $\eta_p^2 = .023$. Paired comparisons with Bonferroni corrections showed that cognate items were recognized significantly more quickly than FC and control items, which did not differ from each other (see Table 8). There was no effect of LOI, $F(1, 64.65) = 0.1$, $p = .75$, $\eta_p^2 = .000$, and no interactions. Due to our theoretical interest, we examined performance separately for each LOI group and each word type. These analyses revealed that the word type effect manifested similarly in both the Hebrew and the English LOI groups (see Table 8), and that there was no effect of LOI in any of the word types (see Appendix S5 in the Supporting Information online). The RT

Table 8 Summary of beta estimates of log reaction times for pairwise comparisons with Bonferroni corrections for multiple comparisons for the word type effect in the translation recognition test within each language of instruction (LOI) group separately

Group	Comparison	*b*	*SE*	95% CI	χ^2	*p*
Across LOI	Cognates vs. controls	0.27	0.03	[0.21, 0.34]	103.85	<.001
	FCs vs. controls	0.04	0.03	[−0.02, 0.10]	2.45	.352
	Cognates vs. FCs	−0.24	0.03	[−0.31, −0.16]	64.94	<.001
Hebrew	Cognates vs. controls	0.26	0.03	[0.18, 0.33]	73.18	<.001
	FCs vs. controls	0.04	0.03	[−0.03, 0.11]	2.11	.438
	Cognates vs. FCs	−0.21	0.03	[−0.30, −0.14]	46.23	<.001
English	Cognates vs. controls	0.30	0.03	[0.22, 0.37]	94.30	<.001
	FCs vs. controls	0.04	0.03	[−0.03, 0.11]	2.03	.464
	Cognates vs. FCs	−0.26	0.03	[−0.34, −0.18]	62.41	<.001

Note. FCs = false cognates. Three multiple comparisons were used to calculate the Bonferroni adjustment.

increase from Session 1 to Session 2 reached significance for the more difficult FC and control items but not for cognates.

Translation Production

The error analysis revealed a significant effect of LOI, $F(1) = 8.90, p = .003$, such that learning through the L1 Hebrew was less error-prone than learning through the L2 English (see Table 9). In addition, there was a main effect of word type, $F(2) = 54.91, p < .001$. Paired comparisons with Bonferroni corrections revealed that cognates elicited significantly fewer errors than FCs and controls, which did not differ from each other (see Table 10). Although the interaction between LOI and word type was not significant, the pattern of means suggested differential effects (see Figure 1). Due to our theoretical interest, we examined performance in each LOI separately and each word type separately (see Appendix S5 in the Supporting Information online). These analyses revealed that, in the English instruction condition, FC items elicited fewer errors than control items, $b = -0.79$, $SE = 0.38$, 95% CI [−1.54, −0.05], $z = -2.08$, $p = .038$, although this difference did not survive correction for multiple comparisons (see Table 10). In the Hebrew instruction condition this difference did not reach significance (see Table 10). In addition, there was no effect of LOI for cognates, but a significant advantage for learning through L1 Hebrew relative to learning through L2 English for FCs, $F(1) = 6.91, p = .009$, and for control items, $F(1) = 10.66, p = .001$ (see Appendix S5 in the Supporting Information online).

Table 9 Model summary predicting error rate in the translation production test as a function of session, language of instruction (LOI), and word type

Fixed effects	*b*	*SE*	95% CI	*z*	*p*
Intercept	1.48	0.31	[0.87, 2.10]	4.75	<.001
LOI (Hebrew)	−0.89	0.30	[−1.48, −0.31]	−2.98	.003
Word type (cognate)	−4.16	0.41	[−4.96, −3.37]	−10.27	<.001
Word type (FC)	−0.69	0.38	[−1.43, 0.05]	−1.82	.069

Random effects	Variance	*SD*
Participant (intercept)	1.12	1.08
Word type (cognate)	–	–
Word type (FC)	–	–
Session (2nd)	–	–
Item (intercept)	1.23	1.11
LOI (Hebrew)	–	–
Session (2nd)	–	–
Residual		
$R^2_{\text{marginal}} = .37$; $R^2_{\text{conditionnal}} = .64$		

Note. Fixed effects reflect simple effects relative to the reference level when other variables are at their reference level without correction for multiple comparisons. For main effects, see *F* values in the text. FC = false cognate.

Table 10 Summary of odds ratios of error rates for pairwise comparisons with Bonferroni corrections for multiple comparisons for the word type effect in the translation production test within each language of instruction (LOI) group separately

Group	Comparison	*OR*	*SE*	95% CI	*z* ratio	*p*
Across LOI	Cognates vs. controls	64.35	26.08	[24.38, 169.80]	10.28	<.001
	FCs vs. controls	1.99	0.75	[0.80, 4.92]	1.82	.210
	Cognates vs. FCs	0.03	0.01	[0.01, 0.09]	−7.80	<.001
Hebrew	Cognates vs. controls	52.61	24.46	[17.29, 160.09]	8.53	<.001
	FCs vs. controls	1.84	0.75	[0.69, 4.90]	1.49	.412
	Cognates vs. FCs	0.03	0.02	[0.01, 0.12]	−6.63	<.001
English	Cognates vs. controls	73.84	31.14	[26.91, 202.66]	10.20	<.001
	FCs vs. controls	2.21	0.85	[0.89, 5.53]	2.08	0.11
	Cognates vs. FCs	0.03	0.01	[0.01, 0.09]	−7.67	<.001

Note. FCs = false cognates. Three multiple comparisons were used to calculate the Bonferroni adjustment.

Table 11 Model summary predicting log reaction time in the translation production test as a function of session, language of instruction (LOI), and word type

Fixed effects	*b*	*SE*	95% CI	*df*	*t*	*p*
Intercept	3.21	0.02	[3.18, 3.25]	75.89	171.67	<.001
LOI (Hebrew)	–	–	–	–	–	–
Word type (cognate)	−0.19	0.02	[−0.22, −0.15]	55.86	−11.26	<.001
Word type (FC)	−0.02	0.02	[−0.05, 0.02]	53.28	−0.87	.389

Random effects	Variance	*SD*	*r*	*r*
Participant (intercept)	0.01	0.12		
Word type (cognate)	0.00	0.06	−.91	
Word type (FC)	0.00	0.05	−.59	.59
Session (2nd)	–	–	–	–
Item (intercept)	0.00	0.04		
LOI (Hebrew)	–	–		
Session (2nd)	–	–		
Residual	0.02	0.13		
$R^2_{\text{marginal}} = .23$; $R^2_{\text{conditionnal}} = .51$				

Note. Fixed effects reflect simple effects relative to the reference level when other variables are at their reference level without correction for multiple comparisons. For main effects, see *F* values in the text. FC = false cognate. *rs* reflect the correlation matrix among the random effects.

In the RT analyses, there was a significant effect of word type, $F(2, 51.20) = 76.08$, $p < .001$, $\eta_p^2 = .095$ (see Table 11). Paired comparisons with Bonferroni correction showed that cognate items were recognized significantly more quickly than FC and control items, which did not differ from each other (see Table 12). There was no effect of LOI and no interactions as these variables were not maintained in the selected model. We nonetheless examined performance separately for each LOI and each word type, which revealed that the word type effect manifested similarly in both the Hebrew and the English instruction conditions (see Table 12) and that there was no effect of LOI for any of the word types (see Appendix S5 in the Supporting Information online).

Discussion

The goal of the current study was to examine whether learning of novel words is better when learning is conducted through the L1 versus the L2 of bilingual speakers. L1 Hebrew–L2 English bilingual participants learned three types of

Table 12 Summary of beta estimates of log reaction times for pairwise comparisons with Bonferroni corrections for multiple comparisons for the word type effect in the translation production test within each language of instruction (LOI) group separately

Group	Comparison	*b*	*SE*	95% CI	χ^2	*p*
Across LOI	Cognates vs. controls	0.19	0.02	[0.14, 0.23]	126.81	<.001
	FCs vs. controls	0.02	0.02	[−0.03, 0.06]	0.76	1.00
	Cognates vs. FCs	−0.17	0.02	[−0.22, −0.13]	94.89	<.001
Hebrew	Cognates vs. controls	0.19	0.02	[0.14, 0.25]	79.27	<.001
	FCs vs. controls	0.02	0.02	[−0.04, 0.07]	0.75	1.00
	Cognates vs. FCs	−0.17	0.02	[−0.23, −0.12]	59.68	<.001
English	Cognates vs. controls	0.18	0.02	[0.12, 0.23]	76.53	<.001
	FCs vs. controls	0.01	0.01	[−0.05, 0.08]	0.26	.612
	Cognates vs. FCs	−0.16	0.02	[−0.21, −0.11]	67.77	<.001

Note. FCs = false cognates. Three multiple comparisons were used to calculate the Bonferroni adjustment.

novel words in German (L3), differing in their degree of lexical-form and semantic overlap with English (cognates, FCs, and control items). Half of the participants learned the novel words through their L1 Hebrew, whereas the other half learned the same set of words through their L2 English, which is also the more typologically similar language to German. Despite this overlap between the L2 and the L3 being learnt, results showed an overall advantage for the group that learned through Hebrew (the L1). An additional finding was that L2 –L3 cognates were learned better across all measures, regardless of the LOI. Finally, there was a FC advantage over control words in the error rate when learning through the more similar L2. In the following paragraphs we address each of these findings.

The Effect of Language of Instruction

In the present study we examined the role of LOI and observed better learning through the L1. Error rates in both the translation recognition and the translation production (from the L1/L2 to the L3) tests were lower for participants who learned the novel German vocabulary via their L1 (Hebrew) and not via their L2 (English). These findings extend the results of Bogulski et al. (2019), who found that whereas English–Spanish bilinguals outperformed English monolinguals when learning novel Dutch words via their L1 English, this advantage was not found for Spanish–English and Chinese–English bilinguals learning via their L2 English. Thus, in their study a L1 LOI advantage was observed when comparisons were made among different bilingual populations. The current study directly compares two groups sampled from the same bilingual population and showed that learning through the L1 was superior to learning through the L2. To understand this L1 advantage, one may resort to the language regulation mechanisms proposed by Bogulski et al. (2019). Because bilinguals inhibit their L1 while learning and using their L2 (Green, 1998), they accumulate experience in L1 inhibition but not in L2 inhibition. As a result, participants who learned German words through their L1 were more experienced in inhibiting the language through which learning took place, compared to participants learning through their L2. This experience in language regulation may have allowed for better learning through the L1.

Another possible explanation for the advantage in learning through the L1 centers on the differential proficiency patterns in the two languages. When the LOI is also the more proficient language (i.e., the L1), the learner may be better able to process new information due to the higher availability of cognitive resources (Segalowitz, 2003; Tzelgov & Kadosh, 2009). This is especially true when learning more difficult items (control items), because these require learning of a novel form and mapping it to a meaning (Tokowicz & Degani, 2015). In the current data, there is suggestive evidence that the effects of LOI varied by the difficulty of the to-be-learned materials. Specifically, although the interaction between LOI and word type did not reach statistical significance, follow-up tests revealed that when the learning material was relatively easy (i.e., cognates), LOI did not tend to moderate learning. Conversely, for FC items, which were moderately difficult to learn, a LOI effect (i.e., benefits of L1 over L2 vocabulary instruction) emerged in the production test and not the recognition test. Because in the recognition test the phonological form of the novel German word was auditorily presented, this was an easier task than the production test, in which participants were required to retrieve the phonological form of the German word (e.g., Kaushanskaya & Yoo, 2011). Together, the descriptive findings that the effect of LOI was stronger for more difficult items and in more difficult tests imply that the advantage in learning through the L1 may be at least partially driven by the available cognitive resources that learners have during both the learning task and the test. Learning through the more proficient L1 frees up more cognitive resources to deal with more difficult learning conditions, and thus the L1 LOI advantage tended to be stronger when learning task and test requirements were high. Accordingly, when learning difficult items (control and FC items), and when required to retrieve the phonological form of the novel word (production but not recognition), an advantage for learning through the (dominant) L1 was more prominent.

The language regulation and the proficiency/resource-availability accounts are not mutually exclusive. Instead, it is likely that both experience with regulation of the L1 *and* the availability of more cognitive resources when learning through the more proficient L1 played a role in the observed effects. The joint contribution of both mechanisms is especially likely in the current population, where participants' L1 was also their dominant language. Notably, both accounts predict that learning would be better through the dominant language, even when it is not the first acquired one. Indeed, based on the post hoc analyses they performed, Bogulski et al. (2019) suggested that language regulation depends on language dominance, such that irrespective of the order of acquisition, bilinguals should be better able to inhibit their dominant language and thus use it more effectively as a LOI. Thus, even when bilinguals' dominance profile has changed, such that their first-acquired language has become their less dominant one (Martin et al., 2020), learning of novel words is expected to be better through their dominant, more proficient L2 under both the language regulation and the proficiency/resource-availability accounts. Future studies will test if these predictions are borne out.

The present study thus suggests an advantage for learning novel words through the (dominant) L1. Although follow-up tests suggested that under a certain condition (i.e., when learning cognates) the LOI effect was not present (as found by Tomoschuk, 2019; Tomoschuk et al., 2021), critically, the effect did not reverse: Learning through the L2 was never better than learning through the L1. Thus, although both the L2 and L3 are nonnative languages, and consequently the L2 is more similar to the L3 in terms of learning circumstances, age of onset, memory representations, and in the current case even typological similarity, there was no advantage for learning through the L2.

Of note, the effect of LOI emerged in both the recognition and the production tests, although the two tasks differ in their level of difficulty (see also, e.g., Degani & Goldberg, 2019; Degani & Tokowicz, 2010a), as well as in the type of word knowledge and associated mechanisms that they tap. Specifically, the recognition test taps the mapping of the novel form to available representations, whereas the production test requires retrieval of the novel form itself. The current findings thus suggest that the LOI affects both of these components of the word-learning process. More prominent effects were apparent in the error-rate data compared to the latency data, probably because RT modulations tend to surface only when performance is accurate enough (for discussion, see Bruyer & Brysbaert, 2011).

The Effect of Word Type

Consistent with previous research, we found a learning advantage for cognate words (e.g., De Groot & Keijzer, 2000). Cognates were learned better across both LOIs, such that learners were able to utilize the overlap in form and in meaning between one of their languages (their L2) and their novel language, irrespective of whether the similar representation was overtly presented during learning (that is, in the Hebrew LOI condition, the English translations were not overtly presented, and still the cognates with English were learned better). This finding is consistent with the suggestion of Bartolotti and Marian (2017), who found that form overlap with either one of a bilingual learner's known languages facilitated learning. In the current study, we did not observe any benefit for learning cognates through the similar language or through the dominant language, even though words (and not pictures, which allow learners to rely on either language as in Bartolotti & Marian, 2017) were used. This pattern probably occurred because cognates were easily learned.

As for FC items, the pattern of means suggested some modulations by LOI, although this interaction did not reach statistical significance, perhaps due to reduced power. Nonetheless, a nuanced follow-up analysis revealed that, whereas the difference between FC and control items did not reach significance when collapsing across LOIs (though it was marginal in the recognition error rate), there was an advantage for FC over control items when learning occurred through the more similar L2 English. Critically, however, examination of Figure 1 reveals that in both error rates and RTs, performance was always numerically better on FC than control items, across both LOIs and in both types of task (recognition vs. production).

FCs were expected to entail two opposing mechanisms (Elias & Degani, under review; Fang, Perfetti, & Stafura, 2017). Form overlap may serve to facilitate learning (as found by Mulík et al., 2019), whereas meaning competition may hinder performance (as found by Rodd et al., 2012). Because FCs were learned better than control items, our findings suggest a greater role for the form facilitation mechanism, especially when learning took place through the language that shares the form similarity with the novel word. That is, it may be that learning through English activated the English lexicon more strongly, allowing learners to benefit more from the lexical-form similarity. This increased English availability may free up cognitive resources needed for the tasks of mapping a new meaning onto a known form and negotiating the meaning competition. Notably, the greater reliance on the form component (leading to facilitation) relative to the meaning component (which would have resulted in interference) may be at least partially biased by the particular characteristics of the current learning paradigm. Because the to-be-learned German words were presented with translations (i.e., word forms rather than images, for instance), learners may have been less likely to develop direct links to concepts (Kroll, Michael, & Sankaranarayanan, 1998). Learning paradigms that put a greater emphasis on meaning (for discussion, see Rice & Tokowicz, 2020) may tilt the balance toward meaning competition having a stronger effect in the case of FC learning, relative to facilitation due to form overlap. Nonetheless, in a recent study, Polish speakers learned via images 24 L2 nonwords that could be cognates, FCs, or noncognate controls (relative to their L1 Polish; Marecka et al., 2020). The findings showed a FC advantage over controls in a production test, with no difference in a recognition test. Thus, even when learning was supported by images (thus potentially offering a more direct link to concepts), form facilitation appeared to outweigh the meaning competition associated with FCs during the initial stages of learning.

The observed word-type effects are informative for considering item-based versus whole-language similarity. Our findings show a greater role of item-based similarity, given that no whole-language typological similarity effect was present (learning German through English was not overall better than learning it through Hebrew). At the same time, item-based similarity affected performance, in that cognates were learned better than controls, and FCs were learned better under some conditions. Thus, when examining novel language learning and the effects of LOI, future research should go beyond manipulations of whole-language typological similarity and examine overlap across languages at finer-grained levels of representation.

Limitations and Future Directions

The current study is the first to examine how LOI affects vocabulary learning among multilingual speakers. The results of the current study reveal that learning through translations given in the L1 can be better than learning through translations in the L2. Notably, however, when interpreting these findings, it is important to keep in mind that the power to detect significant effects, and specifically significant interactions between LOI and word type, was limited (Brysbaert, 2020; Brysbaert & Stevens, 2018). The number of items that participants can learn in such a vocabulary paradigm is limited, but future studies in which more participants are tested per group would be especially informative to examine the replicability and stability of the findings. Nonetheless, by observing significant effects of LOI across measures and sessions, as well as consistent word-type effects, this study sets the ground for future replications and extensions in which additional populations, stimuli, and language domains are tested. Below we set out several areas for possible extension.

First, all participants in the current study were immersed in an environment in which the L1 is widely spoken. Thus, the larger language context in which the study was conducted is biased toward the L1. Given that language context has been shown to affect the level of activation of the two languages among bilingual speakers (Kreiner & Degani, 2015; Linck et al., 2009), future studies should examine whether the LOI effect observed here extends to other environmental language contexts.

Second, some research indicates that bilinguals' lifelong patterns of language use (in terms of interactional context) affect the interplay between the two languages of the speakers and the language control system (e.g., Green & Abutalebi, 2013). The Hebrew–English bilinguals tested in this study tended to use each language in separated life contexts with few instances of language switching (Beatty-Martínez et al., 2020). It is thus unknown whether the LOI effect observed here would manifest similarly in other bilingual populations that differ in interactional domains of language use.

Third, in the current study item-based similarity was operationalized by categorically differentiating high phonological overlap (cognates and FCs) versus low phonological overlap (control items). However, there was variability in the degree to which items overlapped across English and German. Given the importance of phonological overlap underscored by the current findings, future studies in which form overlap is examined as a continuous and complex variable, including both phonological and orthographic overlap, may be especially revealing in the case of word learning, given that this approach has been productive with respect to bilingual word processing (e.g., Dijkstra, Grainger, & Van Heuven, 1999; for discussion, see Degani & Tokowicz, 2010b).

Finally, in addition to the effects of LOI and item-based language similarity exemplified in the current study, novel word learning may also be affected by indirect influences from other cognitive abilities (Hirosh & Degani, 2018). For instance, participants' phonological short-term memory has been shown to positively correlate with word learning (Degani & Goldberg, 2019; Martin & Ellis, 2012) and language learning more generally (e.g., Linck et al., 2013). Similarly, cognitive control abilities, including working memory and inhibitory control (as proposed by Miyake & Friedman, 2012), have been linked to language learning among children (e.g., Gandolfi & Viterbori, 2020), young adults (Linck & Weiss, 2015; Martin & Ellis, 2012, for working memory), and older adults (Pot, Porkert, & Keijzer, 2019). Most relevant to the focus of the current study, to the extent that the ability to regulate irrelevant linguistic knowledge is linked to domain-general cognitive control (e.g., Bartolotti, Marian, Schroeder, & Shook, 2011), and that word learning depends on language regulation (Bogulski et al., 2019), then individuals with better cognitive control are expected to show better learning, especially in the case of FCs when inhibition of the known meaning is beneficial. The low reliability of the cognitive measures collected here, as well as limitations in power, preclude systematic examination of these issues within the current study. These interesting potential moderators therefore await future research.

Conclusion

To conclude, the present study is the first to systematically examine the effects of LOI on novel word learning within the same bilingual population (but see Tomoschuk, 2019 and Tomoschuk et al., 2021, for effects on inhibitory control). The findings show better learning through the more proficient language (the L1). This pattern was more pronounced for difficult-to-learn items and in tests that required production of the novel words. Further, this L1 instruction advantage was present even though the L2 in the current study was typologically similar to the novel learned language. By manipulating item-based similarity, the present study demonstrated that novel language learning is more strongly affected by overlap of specific representations than by whole-language similarity. These findings carry potential educational implications, in that learning a novel language through the less proficient L2 is expected to be less efficient than learning through the L1, at least when the L1 is still the dominant language (i.e., when no switched dominance has occurred). To the extent that learners have the option to choose the language through which to acquire an additional language, the current findings favor the dominant L1.

Final revised version accepted 25 February 2021

Open Research Badges

This article has earned Open Data and Open Materials badges for making publicly available the digitally-shareable data and the components of the research methods needed to reproduce the reported procedure and results. All data and materials that the authors have used and have the right to share are available at http://www.iris-database.org and https://osf.io/habuz/?view_only=bde3ec16027c4671a13e264913ee3d99. All proprietary materials have been precisely identified in the manuscript.

References

Bardel, C., & Falk, Y. (2012). Behind the L2 status factor: A neurolinguistic framework for L3 research. In J. Cabrelli Amaro & J. Rothman (Eds.), *Third language acquisition in adulthood* (pp. 61–78). Amsterdam, The Netherlands: John Benjamins.

Bartolotti, J., & Marian, V. (2017). Bilinguals' existing languages benefit vocabulary learning in a third language. *Language Learning*, *67*(1), 110–140. https://doi.org/10.1111/lang.12200

Bartolotti, J., Marian, V., Schroeder, S. R., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in Psychology*, *2*, 324. https://doi.org/10.3389/fpsyg.2011.00324

Bartoń, K. (2020). MuMIn: multi-Model Inference [Computer software]. R package version 1.43.17. https://CRAN.R-project.org/package=MuMIn

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://www.jstatsoft.org/article/view/v067i01

Beatty-Martínez, A. L., Navarro-Torres, C. A., Dussias, P. E., Bajo, M. T., Guzzardo Tamargo, R. E., & Kroll, J. F. (2020). Interactional context mediates the consequences of bilingualism for language and cognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(6), 1022–1047. https://doi.org/10.1037/xlm0000770

Birdsong, D. (2005). Interpreting age effects in second language acquisition. In J. F. Kroll and A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 109–127). New York, NY: Oxford University Press.

Bogulski, C. A., Bice, K., & Kroll, J. F. (2019). Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. *Bilingualism: Language and Cognition*, *22*(5), 1052–1067. https://doi.org/10.1017/S1366728918000858

Bruyer, R., & Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: Is the Inverse Efficiency Score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychologica Belgica*, *51*, 5–13. https://doi.org/10.5334/pb-51-1-5

Brysbaert, M. (2020). Power considerations in bilingualism research: Time to step up our game. *Bilingualism: Language and Cognition*, 1–6. http://doi.org/10.1017/S1366728920000437

Brysbaert, M., & Biemiller, A. (2017). Test-based age-of-acquisition norms for 44 thousand English word meanings. *Behavior Research Methods*, *49*(4), 1520–1523. https://doi.org/10.3758/s13428-016-0811-4

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: A tutorial. *Journal of Cognition*, *1*(1), 9. http://doi.org/10.5334/joc.10v

Chang, W. (*n.d.*) Plotting means and error bars (ggplot2). Retrieved from http://www.cookbook-r.com/Graphs/Plotting_means_and_error_bars_(ggplot2

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

Degani, T., & Goldberg, M. (2019). How individual differences affect learning of translation-ambiguous vocabulary. *Language Learning*, *69*(3), 600–651. https://doi.org/10.1111/lang.12344

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

Degani, T., & Tokowicz, N. (2010b). Semantic ambiguity within and across languages: An integrative review. *The Quarterly Journal of Experimental Psychology*, *63*, 1266–1303. https://doi.org/10.1080/17470210903377372

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

De Groot, A. M. B., & 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*, 1–56. https://doi.org/10.1111/0023-8333.00110

De Rosario-Martinez, H. (2015). phia: Post-Hoc Interaction Analysis [Computer software]. R package version 0.2-1. https://CRAN.R-project.org/package=phia

Dijkstra, T., Grainger, J., & Van Heuven, W. J. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, *41*, 496–518. https://doi.org/10.1006/jmla.1999.2654

E-Prime 2 [Computer software]. (2019). Pittsburgh, PA: Psychology Software Tools, Inc.

Elias, M., & Degani, T. (under review). Cross-language interactions during novel word learning: The contribution of form similarity and participants' characteristics.

Fang, X., & Perfetti, C. A. (2017). Perturbation of old knowledge precedes integration of new knowledge. *Neuropsychologia*, *99*, 270–278. https://doi.org/10.1016/j.neuropsychologia.2017.03.015

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

Gandolfi, E., & Viterbori, P. (2020). Inhibitory control skills and language acquisition in toddlers and preschool children. *Language Learning*, *70*, 604–642. http://doi.org/10.1111/lang.12388

Ghazi-Saidi, L., & Ansaldo, A. I. (2017). Second language word learning through repetition and imitation: Functional networks as a function of learning phase and language distance. *Frontiers in Human Neuroscience*, *11*, 463. https://doi.org/10.3389/fnhum.2017.00463

Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, *1*(2), 67–81. https://doi.org/10.1017/S1366728998000133

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*, 515–530. https://doi.org/10.1080/20445911.2013.796377

Guo, T., Liu, H., Misra, M., & Kroll, J. F. (2011). Local and global inhibition in bilingual word production: fMRI evidence from Chinese–English bilinguals. *NeuroImage*, *56*, 2300–2309. https://doi.org/10.1016/j.neuroimage.2011.03.049

Hammarström, H., Forkel, R., Haspelmath, M., & Bank, S. (2020). Glottolog 4.2.1. Jena, Germany: Max Planck Institute for the Science of Human History. https://glottolog.org

Hawkins, J. A. (1986). *A comparative typology of English and German*. London, UK: Routledge. https://doi.org/10.4324/9781315687964

Hernández, M., Costa, A., Fuentes, L. J., Vivas, A. B., & Sebastián-Gallés, N. (2010). The impact of bilingualism on the executive control and orienting networks of attention. *Bilingualism: Language and Cognition*, *13*, 315–325. https://doi.org/10.1017/S1366728909990010

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

Hirosh, Z., & Degani, T. (2021a). *Language history questionnaire. Materials from "Novel word learning among bilinguals can be better through the (dominant) first language than through the second language"* [Text/Questionnaire]. IRIS Database, University of York, UK. https://doi.org/10.48316/mxfd-f452

Hirosh, Z., & Degani, T. (2021b). *Experimental instructions and stimuli. Materials from "Novel word learning among bilinguals can be better through the (dominant) first language than through the second language"* [Collection/Instructions and stimuli]. IRIS Database, University of York, UK. https://doi.org/10.48316/3g31-3995

Hirosh, Z., & Degani, T. (2021c). *Vocabulary test. Materials from "Novel word learning among bilinguals can be better through the (dominant) first language than through the second language"* [Language test]. IRIS Database, University of York, UK. https://doi.org/10.48316/rj3k-4r38

Hirosh, Z., & Degani, T. (2021d). *Datasets from "Novel word learning among bilinguals can be better through the (dominant) first language than through the second language"* [Dataset]. IRIS Database, University of York, UK. https://doi.org/10.48316/0k65-7p84

Hirosh, Z., & Degani, T. (2021e). *E-prime scripts for tasks. Materials from "Novel word learning among bilinguals can be better through the (dominant) first language than through the second language"* [Software/Experimental script]. IRIS Database, University of York, UK. https://doi.org/10.48316/p46g-as56

Hirosh, Z., & Degani, T. (2021f). *R code. Materials from "Novel word learning among bilinguals can be better through the (dominant) first language than through the second language"* [Software/Analysis code]. IRIS Database, University of York, UK. https://doi.org/10.48316/j56r-ph15

Hughes, M. M., Linck, J. A., Bowles, A. R., Koeth, J. T., & Bunting, M. F. (2014). Alternatives to switch-cost scoring in the task-switching paradigm: Their reliability and increased validity. *Behavior Research Methods*, *46*, 702–721. https://doi.org/10.3758/s13428-013-0411-5

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

Kaushanskaya, M., & Marian, V. (2009a). Bilingualism reduces native-language interference during novel-word learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 829–835. https://doi.org/10.1037/a0015275

Kaushanskaya, M., & Marian, V. (2009b). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*, 705–710. https://doi.org/10.3758/PBR.16.4.705

Kaushanskaya, M., & Yoo, J. (2011). Rehearsal effects in adult word learning. *Language and Cognitive Processes*, *26*(1), 121–148. https://doi.org/10.1080/01690965.2010.486579

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

Kemp, L. S. & McDonald, J. L. (2021). Second language vocabulary acquisition: The effects of semantic relatedness, form similarity, and translation direction. *Language Learning*. https://doi.org/10.1111/lang.12449

Kilgarriff, A., Baisa, V., Bušta, J., Jakubícˇek, M., Kovárˇ, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (2014). The Sketch Engine: Ten years on. *Lexicography*, *1*(1), 7–36. https://doi.org/10.1007/s40607-014-0009-9.

Kreiner, H., & Degani, T. (2015). Tip-of-the-tongue in a second language: The effects of brief first-language exposure and long-term use. *Cognition*, *137*, 106–114. https://doi.org/10.1016/j.cognition.2014.12.011

Kroll, J. F., Michael, E., & Sankaranarayanan, A. (1998). A model of bilingual representation and its implications for second language acquisition. In A. F. Healy & L. E. Bourne Jr. (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 365–395). Mahwah, NJ: Lawrence Erlbaum.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

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

Linck, J. A., Hughes, M. M., Campbell, S. G., Silbert, N. H., Tare, M., Jackson, S. R., Smith, B. K., Bunting, M. F. and Doughty, C. J. (2013). Hi-LAB: A new measure of aptitude for high-level language proficiency. *Language Learning*, *63*, 530–566. https://doi.org/10.1111/lang.12011

Linck, J. A., Kroll, J. F., & 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. https://doi.org/10.1111/j.1467-9280.2009.02480.x

Linck, J. A., & Weiss, D. J. (2015). Can working memory and inhibitory control predict second language learning in the classroom? *Sage Open*, *5*(4). http://doi.org/10.1177/2158244015607352

Lotto, L., & de Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning*, *48*(1), 31–69. https://doi.org/10.1111/1467-9922.00032

Lüdecke, D. (2021). sjPlot: Data Visualization for Statistics in Social Sciences [Computer software]. R package version 2.8-7. https://CRAN.R-project.org/package=sjPlot

Lüdecke, D. (2021). sjstats: Statistical Functions for Regression Models [Computer software]. R package version 0.18.1. https://CRAN.R-project.org/package=sjstats.

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

Marian, V., Blumenfeld, H. K., & 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*(4), 940–967. https://doi.org/10.1044/1092-4388(2007/067

Martin, J. M., Altarriba, J., & Kazanas, S. A. (2020). Is it possible to predict which bilingual speakers have switched language dominance? A discriminant analysis. *Journal of Multilingual and Multicultural Development*, *41*(3), 206–218. https://doi.org/10.1080/01434632.2019.1603236

Martin, K. I., & Ellis, N. C. (2012). The roles of phonological short-term memory and working memory in L2 grammar and vocabulary learning. *Studies in Second Language Acquisition*, *34*, 379–413. https://doi.org/10.1017/S0272263112000125

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*(1), 8–14. http://doi.org/10.1177/0963721411429458

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, *4*(2), 61–64. http://doi.org/10.20982/TQMP.04.2.P061

Mulík, S., Carrasco-Ortiz, H., & 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*, 1024–1040. https://doi.org/10.1177/1367006918781061

Otwinowska, A., Foryś-Nogala, M., Kobosko, W., & Szewczyk, J. (2020). Learning orthographic cognates and non-cognates in the classroom: Does awareness of cross-linguistic similarity matter? *Language Learning*, *70*, 685–731. https://doi.org/10.1111/lang.12390

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

Pot, A., Porkert, J., & Keijzer, M. (2019). The bidirectional in bilingual: Cognitive, social and linguistic effects of and on third-age language learning. *Behavioral Sciences*, *9*(9), 98. http://doi/org/10.3390/bs9090098

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

R Core Team. (2019). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Retrieved from www.R-project.org

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

Ringbom, H., & Jarvis, S. (2009). The importance of cross-linguistic similarity in foreign language learning. In M. H. Long & C. J. Doughty (Eds.), *Handbook of language teaching* (pp. 106–118). Chichester, UK: Wiley–Blackwell.

Rodd, J. M., Berriman, R., Landau, M., Lee, T., Ho, C., Gaskell, M. G., & Davis, M. H. (2012). Learning new meanings for old words: Effects of semantic relatedness. *Memory & Cognition*, *40*, 1095–1108. https://doi.org/10.3758/s13421-012-0209-1.

Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research*, *27*, 107–127. https://doi.org/10.1177/0267658310386439

Schepens, J. J., van der Slik, F., & van Hout, R. (2016). L1 and L2 distance effects in learning L3 Dutch. *Language Learning*, *66*, 224–256. https://doi.org/10.1111/lang.12150

Schwartz, M., Geva, E., Share, D. L., & Leikin, M. (2007). Learning to read in English as third language: The cross-linguistic transfer of phonological processing skills. *Written Language and Literacy*, *10*, 25–52. https://doi.org/10.1075/wll.10.1.03sch

Schwartz, M., Kahn-Horwitz, J., & Share, D. L. (2014). Orthographic learning and self-teaching in a bilingual and biliterate context. *Journal of Experimental Child Psychology*, *117*, 45–58. https://doi.org/10.1016/j.jecp.2013.08.008

Segalowitz, N. (2003) Automaticity and second language acquisition. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 382–408). Oxford, UK: Blackwell.

Shatil, E., & Share, D. L. (2003). Cognitive antecedents of early reading ability: A test of the modularity hypothesis. *Journal of Experimental Child Psychology*, *86*(1), 1–31. http://doi.org/10.1016/s0022-0965(03)00106-1

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

Tomoschuk, B. (2019). *The acquisition and mechanisms of lexical regulation in multilinguals* (Unpublished doctoral dissertation). University of California San Diego, San Diego, USA.

Tomoschuk, B, Duyck, W, Hartsuiker, R. J., Ferreira, V. S., & Gollan, T. H. (2021). Language of instruction affects language interference in the third language. *Bilingualism: Language and Cognition*, 1–12. https://doi.org/10.1017/S1366728921000043

Tzelgov, J., & Kadosh, R. C. (2009). From automaticity to control in bilinguals. *Trends in Cognitive Sciences*, *13*(11), 455. https://doi.org/10.1016/j.tics.2009.08.007

Ullman, M. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition*, *4*, 105–122. https://doi.org/10.1017/S1366728901000220

Voeten, C. C. (2019). buildmer: Stepwise Elimination and Term Reordering for Mixed-Effects Regression [Computer software]. R package version, 1. https://CRAN.R-project.org/package=buildmer

Wechsler, D. (1997). *WMS-III: Wechsler Memory Scale administration and scoring manual*. San Antonio, TX: Psychological Corporation.

Westfall, J. (2016). PANGEA: Power ANalysis for GEneral Anova designs. http://jakewestfall.org/publications/pangea.pdf

Wilson, M. D. (1988). The MRC psycholinguistic database: Machine-usable dictionary, version 2. *Behavior Research Methods, Instruments & Computers*, *20*, 6–11. https://doi.org/10.3758/BF03202594

Yoo, J., & Kaushanskaya, M. (2012). Phonological memory in bilinguals and monolinguals. *Memory & Cognition*, *40*, 1314–1330. https://doi.org/10.3758/s13421-012-0237-x

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's website:

Appendix S1. Full Set of Stimuli.
Appendix S2. Norming Study for Stimuli Selection.
Appendix S3. Stimuli Characteristics.
Appendix S4. Selected Models.
Appendix S5. Model Summaries.
Appendix S6. Boxplot Figures for the Effects of Word Type and Language of Instruction.

Appendix: Accessible Summary (also publicly available at https://oasis-database.org)

Should Bilinguals Learn Novel Words With Translations in Their First-Language or in Their Second-Language?

What This Research Was About and Why It Is Important

Bilingual speakers can learn foreign vocabulary through translations in their first-language, or through translations in their second-language, but which is better? To test this, adult native Hebrew speakers, with English as their second-language, learned words in German either with Hebrew (native language) translations, or with English translations (a second-language, that is more similar to German). Some of the learned German words were similar to English in sound and meaning (for instance *Maus* in German translates as "mouse" to English), some were similar to English in sound but differed in meaning (*Gift* in German means "poison" in English), and some were dissimilar (*Saft* in German means "juice"). We found that generally bilinguals learned the words better with translations in their first-language, especially when words were dissimilar across the second and third languages. Similar words (like *Maus*) were the easiest to learn. Similar words with a different meaning (*Gift*) tended to be easier to learn with the English translation. Overall, learning with translations in the native language, in which speakers are more proficient and experienced, resulted in better learning.

What the Researchers Did

- 59 adult native Hebrew speakers, who have learned English in school as their second-language, participated.
- Each participant learned 55 German words, presented via headphones, so learners never saw German spelling.

- Half of the participants learned German words with their Hebrew translations, and half with English translations.
- The German words to be learned could be similar to English in sound and meaning, in sound but not meaning, or dissimilar. All German words did not sound like words in Hebrew.
- After learning, participants were asked to (a) recognize which of four translations (in Hebrew or in English, in line with their learning condition) were the correct translation of the German word, and (b) to say the German word out loud when seeing its translation (in Hebrew or English, as per their learning condition).

What the Researchers Found

- Bilinguals learned the German words (especially dissimilar ones) better with their first-language (Hebrew) translations than their second-language (English) translations, even though English is more similar to German.
- Words that were similar in sound and meaning were easiest to learn, regardless of the LOI.
- Words that were similar in sound but not meaning tended to be easier to learn, especially through English.

Things to Consider

- Learning through the first-language was generally better, perhaps because processing the first-language is easier, freeing up cognitive resources to learn the novel words, and because bilinguals may be more experienced in "ignoring" their first-language in order to process information in other languages.
- To the extent that bilingual learners can choose the language through which to learn an additional language, the current findings favor the more proficient firstly acquired language as the LOI.

Materials, data, open access article: Materials and data are publicly available at: https://osf.io/habuz/?view_only=bde3ec16027c4671a13e264913ee3d99 and iris-database.org.

How to cite this summary: Hirosh, Z., & Degani, T. (2021). Should bilinguals learn novel words with translations in their first-language or in their second-language? *OASIS Summary* of Hirosh & Degani (2021) in *Language Learning*. https://oasis-database.org