LANGUAGE LEARNING
A Journal of Research in Language Studies

Language Learning ISSN 0023-8333

EMPIRICAL STUDY

How Individual Differences Affect Learning of Translation-Ambiguous Vocabulary

Tamar Degani and Miri Goldberg
University of Haifa

This study examined interactions of word and learner characteristics during foreign vocabulary learning, focusing on translation ambiguity and individual differences in cognitive resources and linguistic background (language proficiency, multilingual experience). Fifty-three native Hebrew speakers and Russian–Hebrew multilinguals learned the phonological form of target Arabic words along with their Hebrew translations and definitions. The mapping could be translation ambiguous, with a single Hebrew word translated into two Arabic words (one-to-many) or translation unambiguous (one-to-one mapping). Results from translation production and meaning recognition tests revealed that translation-ambiguous words were more difficult to learn than translation-unambiguous words. This disadvantage did not dissipate with time, and learners' phonological short-term memory was associated with increased translation ambiguity costs. Learners' proficiency in the language through which learning took place (Hebrew), but not degree of multilingualism, modulated learning. Findings underscore the importance of item and learner interactions, clarifying the multilingualism effect in novel language learning.

Keywords vocabulary learning; translation ambiguity; multilingualism; phonological memory; second language

Introduction

With globalization, the need to learn an additional language as an adult is becoming increasingly relevant, and many individuals in the world use more than one language in their daily lives (Cenoz, Hufeisen, & Jessner, 2003;

This work was funded by the *Language Learning* Small Grants Research Program and by EU-FP7 grant CIG-322016 to Tamar Degani. The authors would like to thank Zoya Hirosh and Mariana Elias for their valuable assistance in preparing this manuscript.

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

DOI: 10.1111/lang.12344

Edwards, 2004). One critical dimension of learning another language is learning its vocabulary. Indeed, the quality of word knowledge has been highlighted as a prerequisite to reading comprehension in speakers' first language (L1) and their second language (L2) (e.g., Pasquarella, Gottardo, & Grant, 2012; Perfetti, 2007). Interestingly, not all words are equally easy to learn (e.g., de Groot & van Hell, 2005; Degani & Tokowicz, 2010), and not all learners approach vocabulary learning with similar capacities, which include cognitive resources and linguistic background (Kaushanskaya, 2012; Martin & Ellis, 2012). Moreover, certain words may be more difficult to learn for some learners (see Kaushanskaya, Marian, & Yoo, 2011, for interactions between learners' gender and their familiarity with the phonological makeup of words). Therefore, the goal of this investigation was to test whether the interaction of word characteristics and learner characteristics could explain variability in foreign vocabulary learning. We specifically focused on how background characteristics of learners, such as their memory capacity and linguistic background, influence their learning of difficult-to-learn translation-ambiguous words.

Background Literature

Effects of Word Characteristics in L2 Vocabulary Learning

Previous research has shown that some word types are more difficult to learn than others. For example, learning abstract words is substantially more difficult than learning concrete words (for review, see de Groot & van Hell, 2005), and similarly noncognate translations are more difficult to learn than cognate translations (e.g., Lotto & de Groot, 1998). A recently highlighted dimension of difficulty in word learning is the mapping between translations across languages (Degani & Tokowicz, 2010; Degani, Tseng, & Tokowicz, 2014). In particular, not all translations have a one-to-one unique translation in another language, creating a situation of translation ambiguity. For instance, the ambiguous Hebrew word /makor/ can be translated into Arabic as either /munqa:r/ meaning "beak" or as /masˤdar/ meaning "source." Such translation ambiguity may create challenges for learning because adult learners rely on their fully formed and ingrained L1 system, and the indirect L1–L2 mappings of a translation-ambiguous word requires substantial remapping of existing form-to-meaning connections (Tokowicz, 2014). Specifically, a learner not only needs to associate a foreign language word to an existing meaning and the relevant L1 word—as in the case of one-to-one mapping—but this learner is also required to create a one-to-many mapping between a single form and a single meaning to two novel foreign language words (for a schematic illustration, see Eddington & Tokowicz, 2013, Figure 3, p. 445).

Interestingly, translation ambiguity is a prevalent phenomenon across languages (e.g., Tokowicz, Kroll, de Groot, & van Hell, 2002; for review, see Degani, Prior, Eddington, da Luz Fontes, & Tokowicz, 2016) and appears to impact language processing and representations (e.g., Tokowicz & Degani, 2010). For instance, among proficient bilingual speakers of English and Dutch, Tokowicz et al. (2002) observed that translation-unambiguous English–Dutch pairs (e.g., the English word arrow and its Dutch translation *pijl*, corresponding to a one-to-one mapping) were rated as more similar in meaning to each other than translation-ambiguous pairs (e.g., the English word change and its two Dutch translations *verandering* denoting the result of alteration and *wisselgeld* denoting coins of small denomination, corresponding to a one-to-many mapping). Moreover, Tokowicz and Kroll (2007; see also Boada, Sánchez-Casas, Gavilán, García-Albea, & Tokowicz, 2012) found that proficient bilinguals were slower and less accurate in producing translation-ambiguous words (e.g., change) than translation-unambiguous words (e.g., arrow).

Critically, the effect of translation ambiguity is not limited to proficient bilinguals. Rather, it affects the beginning stages of learning as well. Previous research with native English speakers learning Dutch vocabulary found slower reaction times (RTs) and lower accuracy when learning translation-ambiguous words (Degani & Tokowicz, 2010). Degani and Tokowicz taught native English speakers a set of Dutch words that could map onto English unambiguously (single Dutch word corresponding to a single English translation, such as *pijl* and arrow) or ambiguously (two Dutch words corresponding to a single English translation, such as *verandering* and *wisselgeld* to change). The lexical forms of the to-be-learned Dutch words were presented visually along with the corresponding English translation and a relevant definition over multiple sessions. Translation recognition and production tests revealed a substantial accuracy decrement for translation-ambiguous words that did not disappear with time (up to two weeks following initial learning). Further, larger decrements were observed for translation-ambiguous words that were two Dutch synonyms for the same meaning (e.g., for the English sky, the corresponding synonymous Dutch words *lucht* and *hemel*) than for translation-ambiguous words that were Dutch translations of a meaning-ambiguous English word (e.g., *change*). In a second study, Degani et al. (2014) replicated the translation-ambiguity disadvantage with the same stimuli and further showed that teaching the two alternative translations together led to better learning than teaching them apart.

Degani and Tokowicz (2010) proposed that the ambiguity disadvantage stems from the challenge of mapping two foreign language labels to an undifferentiated concept and highlighted two possible mechanisms for the effect. First, active competition between the different alternative translations may hinder performance, which is in line with assumptions of competitive inhibitory interactions between connected representations in interactive activation models (Doherty, 2004; McClelland & Rumelhart, 1981). Second, translation-ambiguous words may suffer from a decrease in associative strength, similar to the fan effect (e.g., Anderson, 1974). The fan effect assumes that as more facts are associated with a concept, each fact has decreased probability of occurrence and thus lower associative strength. As a result, the more facts that are connected to a concept, the longer individuals take to recall the concept. In the case of translation-ambiguous words, a learner would take longer to retrieve a translation if additional translations are linked to the same word.

The majority of previous studies have examined the one-to-many direction of ambiguity, where a single L1 word was mapped onto two foreign language vocabulary items (for an exception focusing on many-to-one mapping, see Bracken, Degani, Eddington, & Tokowicz, 2017). This was also the focus of our study. In particular, in our study, we examined the generalizability of the translation-ambiguity disadvantage in two important ways. First, in previous studies the to-be-learned materials were presented visually with no phonological input throughout the study, placing the emphasis on orthography (for related discussion of the emphasis on orthography in visual presentation, see Dijkstra, Grainger, & van Heuven, 1999). In contrast, in our study, we aurally presented the foreign language Arabic words in reference to visually presented Hebrew translations. Although both Hebrew and Arabic are Semitic languages, their orthographic representations are completely nonoverlapping (see also Degani, Prior, & Hajajra, 2018). Therefore, with the auditory presentation, we avoided the need for transliteration, placing the emphasis on phonology in learning. This design allowed us to test production of the learned Arabic phonological form based on previous encounters with it during the learning process.

Second, because phonological and not orthographic forms were presented during learning, our study allowed us to test the translation-ambiguity disadvantage in different-script languages. To date, the translation-ambiguity disadvantage has mostly been documented for same-script materials, such that the to-be-learned words used the same orthographic system (e.g., Dutch–English and German–English). Notably, the results of a recent study testing the learning of new meanings of known foreign language words indicated that the translation-ambiguity disadvantage may be present across scripts (Lu, Wu, Dunlap, & Chen, 2017; see also Fang, Perfetti, & Stafura, 2017; Rodd et al., 2012). In particular, Lu et al. (2017) taught native Chinese speakers ambiguous and unambiguous English pseudowords, withholding the presentation of the second meaning of the ambiguous words until the second session (see also Degani et al., 2014). The to-be-learned English pseudowords were presented visually along with their meaning (depicted in Chinese characters). The results from the foreign language to L1 translation production tests administered throughout learning revealed a translation-ambiguity disadvantage in the RT measures. No difference was observed between ambiguous and unambiguous words in accuracy rates, likely due to a ceiling effect. In our study, we directly tested for the translation-ambiguity disadvantage in a different set of dissimilar-script languages, teaching Hebrew speakers novel Arabic words.

Testing dissimilar-script languages is important for several reasons. First, this approach allows complete decoupling of the roles of phonology and orthography during learning so that any observed translation-ambiguity effect can be unequivocally traced to the phonological level. Second, one may argue that the lexical system of different-script multilinguals is organized differently from that of same-script multilinguals because the nonoverlapping orthographies allow for functional separation between the languages (e.g., Goral, 2018; Jiang, 2018; van Heuven & Wen, 2018; but see Degani et al., 2018). Thus, results from same-script multilinguals may not generalize to different-script multilinguals. Indeed, testing different-script languages is an important step in generalizing findings observed with English (Share, 2008) and testing the role of language similarity during learning more generally.

Effects of Learner Characteristics in L2 Vocabulary Learning

An additional source of variance in foreign language learning centers on individual differences between learners. Learners may bring to the learning situation different cognitive and linguistic abilities, including differences in memory capacity, attention, and executive functions as well as in existing vocabulary and other linguistic representations. The focus of our study had two critical dimensions: (a) cognitive resources, specifically memory capacity, and (b) linguistic background.

Cognitive Resources

Learning a novel word auditorily involves learning the sequential sound pattern and its (arbitrary) mapping to meaning. Such a process may depend on a learner's ability to memorize phonological sequences. Support for this conjecture comes from the literature on L1 and L2 development among children, showing positive associations between phonological short-term memory and vocabulary growth (Baddeley, Gathercole, & Papagno, 1998; Gathercole & Baddeley, 1993; Hummel & French, 2010). Further, results of recent research with adult learners have been consistent with this suggestion. In particular, Martin and Ellis (2012) examined the correlation between adult learners' memory capacity and their L2 vocabulary and grammar learning. Of relevance, they observed a significant positive correlation between learners' phonological short-term memory as measured by the nonword repetition and nonword recognition tests and their vocabulary learning scores when words were presented to the learners aurally (Gathercole, Pickering, Hall, & Peaker, 2001). Kaushanskaya (2012) further showed that adult monolingual speakers with a larger phonological short-term memory capacity outperformed monolinguals with a smaller memory capacity when learning novel words aurally. The advantage associated with increased phonological short-term memory was restricted, however, to learning items that were composed of nonnative phonemes but was not present when the to-be-learned items included only familiar phonemes. Thus, enhanced phonological short-term memory capacity appears to contribute to learning of novel vocabulary, especially when the to-be-learned items include unfamiliar phonological sequences. In our study, we tested participants' phonological short-term memory and examined how this ability modulated auditory learning of novel foreign language words. Because the to-be-learned Arabic material included some phonemes that do not exist in the participants' known languages (Saiegh-Haddad & Henkin-Roitfarb, 2014), phonological short-term memory was expected to modulate learning.

Moreover, phonological short-term memory may modulate learners' ability to learn translation-ambiguous words. In particular, Degani and Tokowicz (2010) suggested that translation-ambiguous words may be more difficult to learn due to active competition among the alternative foreign language translations. Accordingly, when learners are asked to retrieve the newly learned foreign language word, the availability of the phonological form of the alternative foreign language translation may hinder their performance, and such competition may be stronger for individuals who are able to retrieve both phonological forms accurately. For instance, when asked to learn the Arabic /munqa:r/ as a translation for the Hebrew /makor/, individuals who are better able to maintain in their phonological short-term memory the alternative correct Arabic translation /masˤdar/ are likely to be more strongly affected by the existence of translation ambiguity compared to individuals who are less able to maintain the alternative translation.

Similarly, phonological short-term memory may modulate the translation-ambiguity effect if reduced associate strength (a fan effect) underlies previously observed translation-ambiguity learning disadvantage. In particular, individuals with increased memory may be better able to create and maintain multiple representations and consequently suffer more from reduced associative strength for translation-ambiguous linked representations compared to individuals who are less successful in establishing multiple associations. Thus, individuals with increased phonological short-term memory are likely to show a larger translation-ambiguity effect. Accordingly, we tested whether individuals' phonological short-term memory modulated overall learning, as well as specifically learning of translation-ambiguous words.

An additional memory component implicated in learning additional languages is participants' working memory span. According to the well-known working memory model of Baddeley (2000, 2003), working memory has several components, including temporary storage (the visuospatial sketchpad and the phonological loop), binding and integration (the episodic buffer), and a central executive component responsible for allocating attentional resources to the other three components. Of relevance, the phonological nonword repetition task mentioned above is thought to tap only the storage component, whereas working memory tasks require both storage and manipulation of the to-be-remembered information (Engle, Tuholski, Laughlin, & Conway, 1999). According to this conceptualization, working memory measures are likely to be correlated with phonological short-term memory tasks because both tap the storage component but may account for additional variability in word learning to the extent that the central executive component and manipulation of information are required (Buac, Gross, & Kaushanskaya, 2016). Indeed, in the word learning study described above, Martin and Ellis (2012) observed that learners' working memory capacity as measured by the listening span task (e.g., Harrington & Sawyer, 1992) accounted for unique variance in vocabulary learning even when they took phonological short-term memory into account.

Moreover, working memory has been specifically linked to ambiguity resolution (e.g., Miyake, Just, & Carpenter, 1994). For instance, in the context of within-language ambiguity using a self-paced sentence reading task, Miyake et al. observed that individuals with increased working memory capacity held the two alternative representations longer than those with lower working memory capacity. Accordingly, these individuals may suffer a greater translation-ambiguity disadvantage. Specifically, individuals with a greater working memory capacity have been shown to be more successful in establishing and activating multiple translations simultaneously compared to individuals with reduced working memory. According to competition accounts, the increased simultaneous activation of multiple translations would make individuals with a greater working memory capacity more prone to competition, resulting in increased translation-ambiguity disadvantage. Similarly, according to a reduced associative strength explanation (a fan effect), the stronger established multiple representations for individuals with increased working memory would make them more susceptible to reduced associations for translation-ambiguous items compared to individuals with weaker representations. However, in one study examining whether working memory is linked to the *processing* of translation ambiguity, such a relation was not observed (Michael, Tokowicz, Degani, & Smith, 2011). Specifically, native English speakers who were intermediate learners of Spanish were asked to translate both translation-ambiguous and translation-unambiguous Spanish words. Results showed that participants with a larger working memory span and a greater ability to ignore irrelevant information (based on performance in the Stroop task) were overall more accurate in their translations, but working memory did not differentially influence processing of translation-ambiguous words. In our study, we tested whether working memory may nonetheless influence the *learning* of translation-ambiguous words. We predicted that working memory would not only have a positive effect on foreign language vocabulary learning in general but would also interact with the translation-ambiguity effect such that a larger disadvantage would be observed for individuals with increased working memory capacity.

We examined two additional cognitive resources. First, we measured participants' auditory statistical learning to explore whether individual differences in this ability predict learning. Researchers have widely theorized on the link between statistical learning abilities and language learning (Erickson & Thiessen, 2015; Siegelman, Bogaerts, Christiansen, & Frost, 2017; Siegelman & Frost, 2015), and prior studies have identified links between visual statistical learning abilities and orthographic/morphological L2 learning (Frost, Siegelman, Narkiss, & Afek, 2013). Because statistical learning abilities in different modalities may reflect somewhat nonoverlapping constructs (Frost, Armstrong, Siegelman, & Christiansen, 2015) and may follow a different developmental trajectory (Raviv & Arnon, 2018), in our study, we set out to explore whether auditory statistical learning predicts learning of aurally presented vocabulary among adult learners. Second, we measured participants' nonverbal intelligence to control for baseline differences among learners. Specifically, because we tested a heterogeneous group of participants, it was important for us to verify that any observed learning effects of other individual difference measures were not confounded by baseline differences in nonverbal intelligence because prior studies have observed correlations between memory measures and intelligence (Engle et al., 1999).

Linguistic Background

A second dimension of individual differences that we tested in our study was learners' prior linguistic experience. This construct was examined in two complementary ways: (a) level of proficiency/skill in participants' dominant language and (b) experience with additional languages, namely, degree of multilingualism. With respect to the level of proficiency in learners' dominant language, there is evidence to suggest an association between linguistic abilities in the L1 and those abilities in the L2. For instance, Prior, Goldina, Shany, Geva, and Katzir (2014) observed a moderate positive correlation between high school students' vocabulary scores in their L1 and their vocabulary scores in their L2. Similarly, phonological awareness in the L1 has been linked to phonological awareness in the L2 (Bialystok, McBride-Chang, & Luk, 2005; Saiegh-Haddad & Geva, 2008), and more generally, L1 reading abilities significantly predict later proficiency and achievement in a host of L2 tests (e.g., Sparks, Patton, Ganschow, & Humbach, 2012).

This association between existing linguistic knowledge and performance on newly learned linguistic materials can be explained by two theoretical frameworks (see also Geva, 2014; Melby-Lervåg & Lervåg, 2011). According to the linguistic interdependence hypothesis (Cummins, 1979), this association reflects the reliance of both existing knowledge (e.g., L1) and the newly learned knowledge (foreign language) on a common proficiency/aptitude construct. The assumption is that linguistic abilities in both languages draw from this mutual central processing system (Cummins, 1991), but what this common aptitude entails is still underspecified (Wen, Biedroń, & Skehan, 2017). A related variant of the interdependence hypothesis, the common underlying cognitive processes proposal (Geva, 2014), underscores the relevance of basic cognitive processes, including working memory, phonological awareness, and rapid automatized naming, as comprising the common proficiency construct. Of relevance to our study, according to this theoretical perspective, individuals with higher linguistic abilities in existing languages should more easily learn foreign language vocabulary compared to learners with lower linguistic abilities.

A different theoretical approach, offered to explain the correlation between L1 linguistic abilities and learning of an additional language, is the typological/contrastive approach (Odlin, 1989). This approach highlights the structural similarity between L1 and the additional language. Accordingly, learners rely on their L1 knowledge when learning an additional language, leading to positive transfer when the to-be-learned language is similar to the L1 and to negative transfer when the two systems do not overlap (see also MacWhinney, 2005). According to this perspective, one would expect to observe a correlation between foreign language vocabulary learning and linguistic abilities in a typologically similar language. In our study, Hebrew speakers learned Arabic vocabulary. Because both languages are Semitic languages and share at least some morphological structures (Shimron, 2003), learners may capitalize on the structural similarity between the languages (Norman, Degani, & Peleg, 2016) to more easily parse and process the newly learned Arabic words. Thus, the contrastive approach predicts a correlation between individuals' Hebrew proficiency and their foreign language (Arabic) vocabulary learning.

Notably, our study allowed us to test differential predictions of the two theories by including learners who were proficient in additional languages (English and Russian). According to the interdependence hypothesis, foreign language learning should correlate with proficiency scores in all existing linguistic abilities, thus not only with Hebrew but also with proficiency in other languages (Russian and English), although these are not typologically similar to the foreign language. In contrast, according to the typological/contrastive perspective, foreign language vocabulary learning should benefit from similar representations but not from less similar representations. Thus, this perspective might predict that increased knowledge of multiple typologically different languages would not be associated with greater foreign language learning.

Aside from the level of proficiency in the L1, prior linguistic experience encompasses the degree of multilingualism of learners. Recent studies have shown that multilingual speakers are better at word learning than monolingual speakers (e.g., Bartolotti & Marian, 2012; Kaushanskaya & Rechtzigel, 2012; Kaushanskaya, Yoo, & Van Hecke, 2013; for review, see Hirosh & Degani, 2018; Tokowicz & Degani, 2015). Researchers have proposed that this multilingual advantage in foreign language vocabulary learning stems from both direct transfer (of available representations and practiced skills) and from indirect contributions of mediating cognitive variables, such as superior attentional control abilities and enhanced phonological short-term memory (Hirosh & Degani, 2018). Based on this literature, learners who have had more experience in language learning are expected to outperform less experienced learners. The general prediction is thus that individuals who are more multilingual should outperform those with less multilingual experience. To test this prediction in our study, we had both native Hebrew speakers and Russian–Hebrew speakers learn foreign language Arabic words. All participants were also proficient in English, and thus the sample consisted of a heterogeneous group with varied prior experience with additional languages. We thus conceptualized multilingualism in our study as a continuous variable, including both the number of languages that speakers knew and their proficiency in those languages.

Moreover, the manipulation of translation ambiguity allowed us to test interactions between this item characteristic and learners' degree of multilingualism. Multilinguals may be better able to learn translation-ambiguous words because they have had more experience in mapping two labels to a shared concept (Hirosh & Degani, 2018; Kaushanskaya, Gross, & Buac, 2014). As a result, they may be better able to negotiate the potential competition between the two translations (e.g., /munqa:r/ and /masˤdar/) or may be better able to handle the reduced associative strength of each translation. We therefore tested whether the translation-ambiguity effect was modulated by the degree of multilingualism of the participants.

A recent study has put forth an important caveat to the prediction that multilingualism would generally be associated with a learning advantage. Specifically, Bogulski, Bice, and Kroll (2018) reported evidence that supported the existence of a multilingual advantage only in cases when foreign language vocabulary is learned through L1 translations but not when it is learned via the L2. In that study, native English speakers were compared to English–Spanish, Spanish–English, and Chinese–English bilinguals. All groups learned Dutch vocabulary through English translations. Results indicated a bilingual advantage only for the bilingual group learning the foreign language through their L1 (i.e., English–Spanish bilinguals). The authors proposed that the bilingual advantage in foreign language learning stemmed from their previous experience with L1 regulation (when learning their L2). They assumed that foreign language learning benefits from inhibition of the language from which learning takes place, and thus only bilinguals who have practiced regulating that language enjoy an advantage. Interestingly, in a post hoc analysis, the authors further showed that increased proficiency and dominance in the language through which learning took place was associated with better foreign language learning. In particular, Spanish–English bilinguals who had become more dominant in their L2 English outperformed Spanish–English bilinguals who were Spanish dominant in learning novel Dutch words in association with English translations. This suggested that a multilingual advantage would be observed when bilinguals learn the foreign language words via their L1 or dominant language but not when they learn through their L2 or less dominant language.

Taking this finding into account, in our study, we tested for the multilingual advantage in word learning only when speakers learned through their dominant language. All participants were immersed in the language through which learning took place (Hebrew), and thus, although for some of our participants this language was the L2, their proficiency profile indicated that this was their dominant language. For this population, we thus predicted a multilingual advantage in word learning. Moreover, because we took into account cognitive resources (Kaushanskaya & Marian, 2009), such a multilingual advantage would not be attributed to differences in learners' cognitive profile.

The Present Study

To summarize, our study examined the joint contribution of word and learner characteristics on foreign language vocabulary learning. We focused on translation ambiguity as the central word characteristic by comparing learning of one-to-many translation-ambiguous Arabic words (e.g., /munqa:r/ and /masˤdar/ meaning "beak" and "source," respectively, to the Hebrew /makor/, capturing those two meanings) to learning of translation-unambiguous Arabic words (e.g., /kursi/ to the Hebrew /kise/ meaning "chair"). We extended previous literature demonstrating a translation-ambiguity disadvantage (e.g., Degani & Tokowicz, 2010; Degani et al., 2014) to test whether such a disadvantage was present when learners were provided with the foreign language phonology via auditory presentation rather than their having to derive it from the presented foreign language orthography. By doing so, we were also able to extend the literature to test different script languages. Further, we tested how vocabulary learning of both types of items was modulated by learner characteristics. Here, our focus was on learners' cognitive resources (mainly phonological short-term memory and working memory) and their prior linguistic experience. In particular, we compared more and less proficient Hebrew speakers and examined the role of multilingualism in foreign language learning.

To this end, we tested a heterogeneous group of Hebrew speakers, including native Hebrew speakers and multilingual Russian–Hebrew speakers, in their learning of Arabic vocabulary. The novel vocabulary was presented aurally along with a visually presented Hebrew translation and definition to allow learning of form and meaning. Learning and testing took place over four sessions spanning four weeks to examine longer-term retention. Further, participants completed a battery of individual difference tasks, including phonological and working memory, auditory statistical learning, nonverbal intelligence as well as fluency tests, and self-reported proficiency and use measures to provide a detailed language profile of themselves.

Method

Participants

From a large university in Israel, we recruited 66 participants with no more than minimal prior knowledge of Arabic for this experiment: 35 native

Hebrew speakers and 31 multilingual Russian–Hebrew speakers. The participants received either class credit or payment for participating in each of the experimental sessions. To reduce dropout rates, participants received extra credit or payment for completing all four sessions. At the end of Session 1, participants completed a detailed language history questionnaire, a modified version of the Language Experience and Proficiency Questionnaire (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007; see Table 1), to tap participants' proficiency in and use of each of the languages that they spoke. Russian–Hebrew speakers had learned Russian first, Hebrew second, and English third, whereas native Hebrew speakers had learned Hebrew first and English second. Across sessions, participants also completed tests of cognitive resources (see Table 1).

We had to exclude data from 13 participants from our analyses. We excluded one Russian–Hebrew speaker and three native Hebrew speakers because they were native speakers of additional languages (i.e., simultaneous bilinguals) and one native Hebrew speaker due to technical problems during task administration. We also randomly excluded one native Hebrew speaker to maintain an equal number of participants across experimental versions. Further, based on the data provided by participants in the language history questionnaire, we examined the pattern of language dominance of the Russian–Hebrew speakers. We calculated language dominance either as the ratio of Russian to Hebrew proficiency (following Bogulski et al., 2018) or as the difference between proficiencies in the two languages divided by their sum (following Tomoschuk, Ferreira, & Gollan, 2018). In both cases, we identified seven Russian–Hebrew participants as dominant in their L1 (Russian), and we thus excluded them from analyses. To summarize, we analyzed data from a final set of 53 participants: 30 native Hebrew speakers and 23 multilingual Russian–Hebrew speakers. All included participants had Hebrew as their dominant language (Bogulski et al., 2018).

Materials

We selected a set of 96 Arabic words as stimuli for our study; 24 of these Arabic words corresponded to 12 ambiguous Hebrew words. Each of these ambiguous Hebrew words translates into two different Arabic words (one-to-many translation ambiguity). In addition, we included 48 Arabic words with a single Hebrew translation to construct the unambiguous control type (one-to-one translation unambiguous). The remaining 24 Arabic words were fillers for the purposes of our study. These included ambiguous Arabic words that could correspond to two Hebrew translations ($k = 12$) or to a single joint translation

Table 1 Descriptive statistics of participants' linguistic and cognitive characteristics

Characteristic	L1 Hebrew ($n = 30$) M (SD)	L1 Russian ($n = 23$) M (SD)	All ($N = 53$) M (SD)
Age (years)	24.90 (2.98)	26.26 (4.64)	25.49 (3.81)
Age began Hebrew (years)	0.00 (0.00)*	7.00 (3.90)*	3.04 (4.33)
No. of languages	2.53 (0.57)*	3.26 (0.54)*	2.85 (0.66)
Hebrew proficiency rating[a] (M)	9.50 (0.55)	9.67 (0.43)	9.58 (0.50)
Russian proficiency rating[a] (M)		6.79 (2.00)	
English proficiency rating[a] (M)	7.58 (1.06)	7.51 (1.33)	7.55 (1.17)
Hebrew use rating[b] (M)	7.47 (1.32)	7.56 (1.14)	7.51 (1.23)
Russian use rating[b] (M)		3.28 (1.96)	
English use rating[b] (M)	6.43 (1.71)	6.80 (1.59)	6.59 (1.65)
Reading in Hebrew[c] (%)	78.86 (20.02)*	63.80 (22.99)*	72.32 (22.45)
Reading in Russian[c] (%)		14.79 (17.80)	
Reading in English[c] (%)	20.76 (19.79)	18.75 (10.72)	19.89 (16.37)
Talking in Hebrew[c] (%)	75.86 (23.66)*	60.81 (25.20)*	69.33 (25.25)
Talking in Russian[c] (%)		20.51 (15.60)	
Talking in English[c] (%)	19.14 (17.84)	13.53 (14.78)	16.71 (16.67)
Exposure to Hebrew[c] (%)	77.13 (12.56)*	60.04 (18.15)*	69.72 (17.34)
Exposure to Russian[c] (%)		22.26 (14.00)	
Exposure to English[c] (%)	20.29 (10.69)*	13.92 (9.65)*	17.52 (10.65)
Previous Arabic proficiency[d]	0.37 (.69)	0.22 (.52)	0.30 (.61)
Previous Arabic use[d]	0.13 (.35)	0.13 (.34)	0.13 (.34)
Phonological short-term memory (0–14)	5.90 (1.58)	5.52 (1.16)	5.74 (1.42)
Working memory (0–21)	12.03 (2.13)	11.70 (2.48)	11.89 (2.27)
Nonverbal IQ (0–18)	10.00 (2.41)	9.83 (3.70)	9.93 (3.01)
Hebrew phonemic fluency (M w/min)	13.58 (2.98)	12.46 (4.03)	13.09 (3.48)
Hebrew semantic fluency (M cat/min)	20.73 (3.68)*	18.42 (3.94)*	19.73 (3.93)
Statistical learning (proportion correct)	0.59 (0.15)	0.57 (0.14)	0.58 (0.14)

Note. *A significant difference between the group in a t test ($p < .05$). [a]Language proficiency reflects the average proficiency ratings in reading, writing, conversation, and speech comprehension rated on a scale of 0 (lowest) to 10 (highest). [b]Language use reflects the average use in reading, writing, conversation, internet, listening, and TV watching rated on a scale from 0 (lowest) to 10 (highest). [c]Self-estimated by participants such that all languages sum to 100%. [d]Self-rated on a scale from 0 (lowest) to 10 (highest).

($k = 12$) encompassing the two meanings of the ambiguous Arabic word (see Table 2 for example stimuli and Appendix S1 in the Supporting Information online for the full list of items).

A highly proficient Arabic–Hebrew bilingual determined the Hebrew translations of the Arabic words. We created definitions by consulting two online dictionaries (milon.walla.co.il and www.milog.co.il) and slightly modified them when we needed to keep the definitions short. We presented unambiguous items with only one definition, whereas we provided a separate definition for each of the two possible translations for the ambiguous type. Appendix S1 shows that the Hebrew words that we selected for the translation-ambiguous type were semantically ambiguous such that the two Arabic translations were relatively unrelated in meaning (i.e., termed meaning ambiguous by Degani & Tokowicz, 2010).

We selected stimuli so that, across translation-ambiguous and translation-unambiguous words, items were matched for word length (number of letters, number of phonemes, and number of syllables) for both Hebrew and Arabic and for Hebrew word frequency based on the heTenTen 2014 corpus via Sketch Engine (Kilgarriff et al., 2014; Kilgarriff, Reddy, Pomikálek, & Avinesh, 2010). We obtained phonological form similarity ratings of the Arabic and Hebrew words in a norming study with 34 additional native Hebrew participants who did not participate in the main study. We asked these participants to rate the degree of similarity between each auditorily presented Arabic word and each phonology of the visually presented Hebrew word by pressing a number between 1 (different) and 5 (similar). We normed the Hebrew words in the translation-ambiguous type twice, once with each Arabic translation. Table 3 provides the means and standard deviations for the word characteristics. Across translation type, there were no significant differences in Arabic and Hebrew length, Hebrew frequency, and phonological similarity ratings (all ps $< .05$).

We divided the 96 Arabic words into three versions, counterbalanced across participants. Each version consisted of a different subset of 64 Arabic words, such that each participant learned 32 Arabic words from the translation-unambiguous type (one-to-one mapping) and 16 Arabic words from the translation-ambiguous type (i.e., eight ambiguous Hebrew words, each corresponding to two Arabic translations). The remaining 16 Arabic words in each version were ambiguous Arabic words with two meanings and served as fillers. Thus, each participant learned 64 Arabic words, of which half were translation unambiguous and half were ambiguous in some way. Of these, we examined only items in the one-to-many direction (one ambiguous Hebrew word corresponding to two Arabic translations) in the study reported here.

Table 2 Example of stimuli and of definitions for each translation type

Translation type	Hebrew word	Hebrew IPA	Arabic translation(s)	Arabic IPA	English meaning	Definition in English	Definition in Hebrew
Unambiguous	1. כיסא	kise	1. كرسي	1. kursi	1. chair	1. Furniture for sitting, for one person	1. רהיט, אביזר לישיבה לאדם
Ambiguous	1. מקור	makor	1. منقار 2. مصدر	1. munqa:r 2. masˤdar	1. beak 2. source	1. A bird's organ, used for eating 2. The place from which something derives	1. איבר בציפורה משמש לאכילה 2. המקום שממנו נובע דבר-מה

Table 3 Characteristics of the stimuli used for the two translation types

	Arabic		Hebrew	
Characteristic	Unambiguous ($k = 48$) *M* (*SD*)	Ambiguous ($k = 24$) *M* (*SD*)	Unambiguous ($k = 48$) *M* (*SD*)	Ambiguous ($k = 12$) *M* (*SD*)
Letters	4.06 (1.00)	4.04 (1.12)	4.02 (0.93)	3.83 (0.82)
Phonemes	5.23 (1.33)	5.46 (1.74)	5.10 (1.19)	4.75 (0.97)
Syllables	2.06 (0.84)	2.25 (0.90)	2.06 (0.52)	1.83 (0.39)
Frequency	–	–	37.47 (73.75)	57.49 (48.65)

Note. Frequency ratings for the Arabic words were not obtained. *M* (*SD*) phonological similarity ratings for unambiguous translations = 2.31 (0.22) and for ambiguous translations = 2.31 (0.25).

Procedure

The experiment consisted of four sessions separated by two days, one week, and two weeks, respectively. Training took place in the first two sessions, and testing took place at the end of each training session to increase learning (e.g., Karpicke & Roediger, 2008) and during Sessions 3 and 4. At the beginning of Session 1 and at the end of Session 4, participants completed a pretest and a posttest semantic relatedness judgment task in Hebrew. They had to judge the semantic relatedness of pairs of phrases instantiating different meanings of an ambiguous Hebrew word (e.g., dog bark–tree bark, Degani & Tokowicz, 2013). The Hebrew word was a translation of a learned Arabic word. Because there were no differences on this test before and after learning, this task is not discussed further. At the end of Session 1, participants completed a detailed language history questionnaire, which was modified from the LEAP-Q (Marian et al., 2007). Spread throughout the four sessions, participants performed six individual difference tests, tapping cognitive resources and linguistic background (see details below). The general procedure is summarized in Figure 1. Across the entire protocol, Arabic words were presented only in an auditory form.

Learning

At the beginning of each training task or learning test, participants had eight practice trials during which the experimenter was still present in the room to confirm that they understood the task. Participants were instructed to learn the Arabic words and their meanings during the training cycles, of which there were

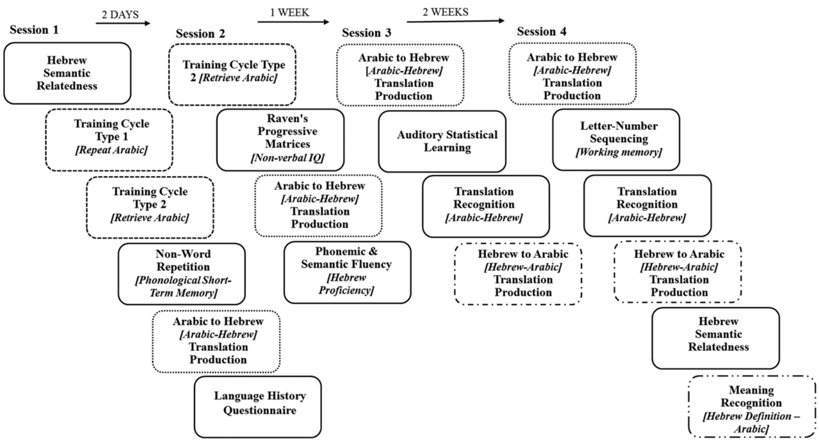

Figure 1 General procedure of training and testing.

two types (following Kang, Gollan, & Pashler, 2013). In the first, which took place only in Session 1, participants were instructed to repeat the Arabic word presented to them (Kaushanskaya & Yoo, 2011). Specifically, in each trial, a fixation cross appeared at the center of the computer screen for 1 second, followed by a blank screen for 500 milliseconds and then by a written Hebrew word with its definition, which appeared for 3 seconds. The word and definition disappeared, and an auditory presentation of a corresponding Arabic word was presented via loudspeakers. Then, a question mark appeared on the screen until the participant's vocal response (repetition of the Arabic word) triggered the voice key. A stimulus onset asynchrony (SOA) of 1,500 milliseconds was included before the beginning of the next trial.

In the second type of learning cycle, which took place in Session 1 and at the beginning of Session 2, the question mark appeared right after the Hebrew word and its definition, and participants were asked to attempt to produce the Arabic word before hearing the correct form. The auditory presentation of the correct Arabic word occurred 1 second after the participant's response. Such retrieval attempts have been shown to strengthen learning (Kang et al., 2013; for review, see Tokowicz & Degani, 2015). An SOA of 1,500 milliseconds was included before the beginning of the next trial.

Each learning cycle included a total of 80 trials: 32 trials of the unambiguous type, 16 trials of the ambiguous type (eight ambiguous Hebrew words, presented twice, once with each of its Arabic translations), and 32 filler trials (16 ambiguous Arabic words, each presented twice, once with each of its definitions). Each Arabic word was presented twice in Session 1 (once in each learning cycle type) and one more time in Session 2, for a total of three repetitions across the entire learning protocol. Two short breaks were introduced into each learning cycle, offering participants the opportunity to rest their eyes and regain focus before continuing to the next trial.

In total, each participant learned 64 Arabic words. The presentation order of translation type was sequential within each learning cycle, such that translation-unambiguous items were presented before translation-ambiguous items, and the order of Arabic words within each type was randomized. During the learning phase, participants' attention was not directly drawn to the fact that some translation-ambiguous words were included in the learning set, however, this was explicitly stated in the translation production tests administered for the first time at the end of Session 1 (see Figure 1).

Testing

We administered four different learning tests aimed at tapping different aspects of learning. The first was an Arabic-to-Hebrew translation production test that we used in all four sessions (see also Degani & Tokowicz, 2010; Degani et al., 2014). This is a backward translation production task, testing participants' memory of the novel Arabic words. This task has been shown to produce an accuracy range that allows for translation-ambiguity effects to emerge (Degani & Tokowicz, 2010; Degani et al., 2014). In the last two sessions, we also tested participants with a Hebrew-to-Arabic translation production test (forward translation production) to tap participants' ability to retrieve the correct phonological form of the Arabic words. Due to our focus on auditory presentation, it was important to verify that learners were able to produce the phonological form of the learned words, although performance on this task was expected to be overall lower (Degani & Tokowicz, 2010). In addition, we administered in the last session a meaning recognition task in which the Arabic word was presented along with a correct or an incorrect definition to tap participants' memory of the link between the Arabic word and its meaning. Finally, we administered a translation recognition test in which the Arabic word was presented along with a correct or an incorrect Hebrew translation; regrettably, correct and incorrect trials were improperly balanced in different translation types, so we have not included the results of this task in the analyses. As in the learning cycles, words were randomly presented within each translation type in each test, but the presentation order of types was blocked (unambiguous→ambiguous→unambiguous→ambiguous, with filler item blocks interspersed). Different unambiguous items were presented before and after the ambiguous items. The instructions of the two translation production tests indicated that translation-ambiguous words would appear twice and instructed learners to provide a different translation word on each presentation. These instructions explicitly pointed out that some words were learned in association with two possible translations.

In the Arabic-to-Hebrew translation production test, participants provided the Hebrew translations of the Arabic words. On each trial, a fixation cross appeared for 1 second followed by an auditory presentation of the Arabic word. Immediately after the offset of the Arabic word, a question mark appeared onscreen until participants' oral response (a Hebrew translation) triggered the voice key. An SOA of 1,500 milliseconds was included before the beginning of the next trial. Participants' responses were recorded for later coding of accuracy. A single researcher coded the responses. Cohen's kappa interrater reliability based on 10% of the data was .96.

In the Hebrew-to-Arabic translation production test, participants provided the Arabic translation upon presentation of the Hebrew word. On each trial, a fixation cross appeared on the computer screen for 1 second followed by a blank screen for 500 milliseconds and then by a Hebrew word until the participants' vocal response (an Arabic word) triggered the voice key. An SOA of 1,500 milliseconds was included before the beginning of the next trial. The Hebrew translations that corresponded to two different Arabic words (ambiguous type) appeared twice, and the participants were instructed to provide a different Arabic word on each presentation. Participants' responses were recorded for later coding for accuracy. Partially correct pronunciations, in which vowels or consonants were omitted or added (e.g., /jibna/ instead of /jubna/), were treated as incorrect trials in these analyses. Slight changes in accent for nonnative phonemes were considered correct. A single researcher coded the responses. Cohen's kappa interrater reliability based on 10% of the data was .75.

Finally, in the meaning recognition test, participants indicated whether a Hebrew definition corresponded to an Arabic word by pressing either the *yes* button or the *no* button with the index finger of their dominant hand. The buttons were placed next to each other so that no hand movement was required. Each trial began with a fixation cross for 1 second followed by a blank screen for 500 milliseconds and then by a written Hebrew definition for 2.5 seconds. Next, an Arabic word was presented aurally, followed by a question mark, until the participants made a response. An SOA of 500 milliseconds was included before the beginning of the next trial. In the translation-unambiguous type, each Arabic word was presented once with its correct definition and once with an incorrect definition (a definition of another Arabic word from the same translation type). Thus, 50% of these pairs (32 pairs) required a correct response. In the translation-ambiguous type, each Arabic word was presented three times, once with its correct definition and twice with an incorrect definition (one unrelated foil and one whose meaning corresponded to the other Arabic translation item of the same ambiguous Hebrew word). Thus, 33% of the ambiguous pairs (16 pairs) required a correct response. Because additional filler items were included, a total of 160 pairs were presented, with 50% requiring a correct response.

Measures of Individual Differences

We tested participants' phonological short-term memory using a Hebrew version of the nonword repetition task (see the syllable span task, Shatil & Share, 2003). In this task, we asked participants to repeat out loud, in the order presented, sets of Hebrew nonwords of two to eight monosyllabic items. The test ended when the participant did not succeed in accurately repeating both sets of a given length. Split-half reliability with Spearman-Brown correction of this test was .67.

We tested working memory span using the Letter–Number Sequencing test (a subtest of the Wechsler Adult Intelligence Scale Version III; Wechsler, 1997) in its Hebrew version (WAIS-III HEB; http://www.psychtech.co.il). The test requires participants not only to maintain information in phonological memory (as is done in the nonword repetition task) but also to perform a mental operation on that information, and it thus taxes working memory capacity. In this task, we presented participants with letter and number strings varying in length from two to eight characters and asked them to repeat each sequence in a different order by first repeating the numbers in ascending order and then repeating the letters in their alphabetic order. Three strings were presented at a given length level. The test ended when participants did not succeed in accurately repeating any of the three strings of a given length. Split-half reliability with Spearman-Brown correction of this test was .73.

We tested statistical learning abilities via an auditory statistical learning test. We conducted the test following the procedures of the auditory-verbal-adjacent task (Siegelman & Frost, 2015). We first asked participants to listen to a 9-minute monologue in a novel language (verbal stimuli) via headphones, including 12 auditory syllable triplets repeated 52 times. They then heard pairs of three-syllable sequences ("words") with a 300-millisecond interval between them and had to indicate by pressing either the 1 or the 2 button which word (first or second) appeared in the language that they had previously heard. Split-half reliability for this test with Spearman-Brown correction was .73. Test–retest reliability of this task has been reported as $r = .63$ (Siegelman & Frost, 2015).

Finally, to control for participants' nonverbal intelligence we employed a short version of the Raven's Progressive Matrices Test (Raven, 1960; adapted from Degani, 2011; Degani & Tokowicz, 2013). In this task, participants had to complete a visual pattern by selecting one of eight alternatives. Reliability of the task has been previously established (Raven & Rust, 2008). Split-half reliability with Spearman-Brown correction of this test was .78.

We collected information about participants' linguistic backgrounds via a language history questionnaire, a modified version of the validated LEAP-Q (Marian et al., 2007). In the questionnaire, participants provided information on their proficiency, use, and learning circumstances of all the languages that they knew and were asked directly whether they had learned Arabic and what their current level of Arabic proficiency and use was.

These subjective proficiency measures were complemented with more objective tasks. Specifically, to measure Hebrew proficiency participants completed the phonemic fluency test and the semantic fluency test (Kavé, 2005). In these tests, participants had to produce in one minute as many words as possible for each of three Hebrew letters (*bet* /b/, *gimel* /g/, and *shin* /š/) and for each of three semantic categories (animals, fruits and vegetables, and vehicles), respectively. Task administration and scoring was identical to that described in Kavé (2005). Based on the current sample, Cronbach's alpha reliability for the phonemic fluency test (three phonemes) was .77, and for the semantic fluency test (three categories) was .77. For the total fluency test, including all six items, Cronbach's alpha reliability was .82.

Data Analysis

We analyzed the data using linear mixed-effects models as implemented in the lme4 (Baayen, Davidson, & Bates, 2008) and lmerTest packages (Kuznetsova, Brockhoff, & Christensen, 2017) in R (Version 3.3.3; R Core Team, 2017). We analyzed accuracy data following a binomial distribution (i.e., logistic linear mixed effects). We analyzed RTs for correct responses and trimmed them to exclude trials with RTs shorter than 200 milliseconds or longer than 4,500 milliseconds, resulting in the exclusion of 5% of the data from the Arabic-to-Hebrew translation production test, 2% from the Hebrew-to-Arabic translation production test, and 7% from the meaning recognition test. We included in the model as fixed effects item characteristics (specifically translation ambiguity) on the one hand and learner characteristics (cognitive resources and linguistic background) on the other, and we examined their interaction with each other.

Table 4 Mean performance (standard error) across tests and sessions for the unambiguous and ambiguous translation types

		Ambiguous	
Test and session	Unambiguous	All trials	No repetitions[a]
Accuracy (% correct):			
Arabic-Hebrew production 1	43 (1)	29 (2)	
Arabic-Hebrew production 2	54 (1)	36 (2)	
Arabic-Hebrew production 3	50 (1)	34 (2)	
Arabic-Hebrew production 4	51 (1)	35 (2)	
Hebrew-Arabic production 3	43 (1)	46 (2)	36 (2)
Hebrew-Arabic production 4	45 (1)	52 (2)	41 (2)
Meaning recognition 4	82 (1)	73 (1)	
Reaction times (ms):			
Arabic-Hebrew production 1	1,045 (24)	1,211 (48)	
Arabic-Hebrew production 2	916 (20)	1,148 (48)	
Arabic-Hebrew production 3	1,053 (26)	1,243 (51)	
Arabic-Hebrew production 4	975 (25)	1,121 (47)	
Hebrew-Arabic production 3	1,336 (25)	1,409 (39)	1,501 (32)
Hebrew-Arabic production 4	1,238 (21)	1,380 (36)	1,508 (32)
Meaning recognition 4	548 (10)	585 (14)	

Note. [a]Means and standard errors for no repetition ambiguous words refer to computations excluding cases in which participants repeated the same Arabic word when presented with the Hebrew ambiguous word for the second time.

Prior to analyses, it was necessary to examine multicollinearity in the learner characteristic predictors.

Results

Table 4 provides mean accuracy and RTs for correct responses by translation type (ambiguous vs. unambiguous) for the final sample of 53 participants. Performance on filler items is not reported here.

Individual Difference Predictors

The study focused on individual differences in cognitive resources (phonological short-term memory, working memory, nonverbal IQ, and auditory statistical learning) as well as in linguistic background, including Hebrew proficiency and multilingual language experience. As a first step, we examined the Pearson correlations among the different cognitive variables (see Table 5). These analyses

Table 5 Pearson correlations among cognitive resources measures

Measure	1	2	3	4
1 Phonological short-term memory	–			
2 Working memory	.50**	–		
3 Nonverbal IQ	–.10	.13	–	
4 Statistical learning	.28*	.18	–.20	–

Note. $^{*}p < .05$, $^{**}p < .01$.

revealed a significant association between phonological short-term memory and working memory and between phonological short-term memory and auditory statistical learning. Thus, before entering these variables into the model predicting learning, we residualized working memory to isolate the portion of working memory variance that was not explained by phonological short-term memory. Similarly, we residualized auditory statistical learning, isolating the portion of auditory statistical learning not explained by phonological short-term memory. There were no other significant correlations among the cognitive resource measures.

We next examined the Pearson correlations among the linguistic background measures, including the objective Hebrew proficiency measures (semantic and phonemic fluency tasks), self-rated ratings of Hebrew proficiency and use (on a 0–10 scale, see Table 1), and in addition, self-rated estimates of the other languages of the speaker. These included the number of spoken languages, self-rated estimates of reading and speaking proficiency (on a 0–10 scale) in the most proficient language other than Hebrew for each participant (which could have been Russian or English for the Russian–Hebrew speakers), and percent of current exposure in reading and speaking to all languages other than Hebrew. Examination of the Pearson correlations (see Table 6) revealed several significant correlations.

To reduce collinearity in the predictors, we applied a principal component analysis to the data. Indeed, the Kaiser-Meyer-Olkin measure of sampling adequacy of .661 and Bartlett's test of sphericity ($p < .001$) indicated that the predictors were highly correlated, suggesting that a principal component analysis was warranted for this dataset. We extracted factors with eigenvalues over 1 and applied a Varimax rotation with Kaiser normalization to increase interpretability of the factors by increasing the likelihood that each original measure would correlate highly with only one factor. Table 7 presents the results of the principal component analysis, including factor loadings, percent of

Table 6 Pearson correlations among linguistic background measures

Measure	1	2	3	4	5	6	7	8	9
1 Hebrew: Phonetic fluency	–								
2 Hebrew: Semantic fluency	.63**	–							
3 Hebrew: Self-rated proficiency	–.05	–.11	–						
4 Hebrew: Self-rated use	–.02	–.13	.20	–					
5 No. of languages	.07	–.24	.14	–.07	–				
6 Other languages: Reading percent	–.12	–.19	–.10	–.20	.49**	–			
7 Other languages: Talking percent	–.01	–.14	–.15	–.01	.50**	.68**	–		
8 Other languages: Exposure percent	–.11	–.27	–.03	–.04	.62**	.61**	.61**	–	
9 Other languages: Self-rated proficiency	–.04	–.15	.48**	.03	.36**	.40**	.35*	.36**	–

Note. $^{*}p < .05$, $^{**}p < .01$.

Table 7 Rotated matrix with the loadings of the different linguistic background variables on three factors based on a principal component analysis

Variables	Multilingualism	Objective Hebrew	Subjective Proficiency
Hebrew: Phonemic fluency	.01	**.91**	–.01
Hebrew: Semantic fluency	–.20	**.87**	–.15
Hebrew: Self-rated proficiency	–.01	.00	**.90**
Hebrew: Self-rated use	–.16	–.14	**.47**
No. of languages	**.77**	.11	.16
Other languages: Reading percent	**.84**	–.10	–.15
Other languages: Talking percent	**.83**	–.02	–.13
Other languages: Exposure percent	**.83**	–.16	–.02
Other languages: Self-rated proficiency	**.55**	.03	**.63**
Unique variance explained (%)	33.87	18.24	16.89
Reliability	.71	.77	.34

Note. Numbers in bold show loadings that are higher than .50. Factor reliability reflects Cronbach's alpha inter-item reliability on the dimensions with loadings above .50 for a given factor.

unique variance explained by each factor, and reliability assessments. Three factors emerged, cumulatively capturing 69% of the variance in the original predictors. Specifically, measures of number of languages and proficiency in, as well as exposure to, languages other than Hebrew, loaded most strongly on the first factor. We thus termed this the Multilingualism factor. Measures of

Table 8 Pearson correlations among individual difference predictors

Measure	1	2	3	4	5	6	7
1 Phonological short-term memory	–						
2 Working memory (residualized)	.00	–					
3 Nonverbal IQ	–.10	.08	–				
4 Statistical learning (residualized)	.00	–.15	–.17	–			
5 Multilingualism factor	–.06	–.05	–.11	.03	–		
6 Objective Hebrew factor	.29*	.30*	–.10	.18	.00	–	
7 Subjective Proficiency factor	–.02	.16	.09	–.07	.00	.00	–

Note. $^{*}p < .05$.

phonemic and semantic Hebrew fluency loaded most strongly on the second factor. We thus referred to this as the Objective Hebrew factor. Finally, self-rated Hebrew proficiency and use as well as self-rated proficiency in the other language loaded most strongly on the third factor, which we termed the Subjective Proficiency factor. Self-rated proficiency in a language other than Hebrew loaded on both the Multilingualism factor and on the Subjective Proficiency factor, likely because these subjective ratings are at least partially explained by individuals' bias in self-report measures (Tomoschuk et al., 2018). Further, the reliability of this factor was rather low. Accordingly, we treated the third factor as most strongly reflecting individual bias in self-report and included it only as a control variable in the model.

Finally, before we entered these individual difference measures into the model predicting learning outcomes, we examined the correlations among these selected predictors (see Table 8). This analysis revealed a correlation between the Objective Hebrew factor and the memory scores, likely reflecting the partial reliance of these memory measures on long-term memory representations (for discussion, see Buac et al., 2016). Critically, all correlations did not exceed .29, alleviating concerns of extreme multicollinearity. Thus, we next examined how these seven individual difference variables and factors contributed to learning. We normalized all measures prior to conducting the analyses.

Model Structure

The main objective of our study was to test whether translation ambiguity impacts auditory vocabulary learning and whether such learning is modulated by individual differences in cognitive resources and linguistic background. To this end, we adopted the following model structure. The effect of translation type was deviation coded as a fixed-effect variable (unambiguous as –0.5 vs.

ambiguous as 0.5). For the translation production tests administered multiple times, we included an additional fixed effect of session, with the first session set as the reference (i.e., Session 1 for the Arabic-to-Hebrew production translation test, and Session 3 for the Hebrew-to-Arabic production translation test), as well as the interaction between translation type and session. In the meaning recognition test, administered only once, session was replaced by response type (*no* vs. *yes* decisions, with *no* decisions set as the reference) to control for potential bias in participants' responses. Further, we included the seven individual difference predictors (phonological short-term memory, residualized working memory, nonverbal intelligence, residualized statistical learning, Multilingualism factor, Objective Hebrew factor, and Subjective Proficiency factor) as fixed effects. Finally, to test relevant modulations of translation ambiguity with theoretically motivated learner characteristics, we included the interactions of translation type with phonological short-term memory, residualized working memory, the Multilingualism factor, and the Objective Hebrew factor. To simultaneously account for variance due to participants and for variance related to items, we included by-participant and by-item intercepts as random effects. Due to our focus on individual differences, we did not include by-participant slopes. The model estimates across measures are summarized in Appendix S2 in the Supporting Information online.

Translation Ambiguity Effect

The main objective of the study was to test the degree to which auditory learning is modulated by translation ambiguity. As shown in Figure 2 (and in Appendix S2), the results revealed a consistent pattern by which translation-unambiguous words were learned better than translation-ambiguous words. The effect was reliable for the Arabic-to-Hebrew translation production test for both accuracy and RT as well as for the accuracy for the meaning recognition test. Interestingly, there was no reliable main effect of translation type for the Hebrew-to-Arabic translation production test, and the pattern of means appeared to suggest an ambiguity advantage (see Figure 2). However, ambiguous words were repeated twice in this task, and participants often provided the same Arabic translation twice (35% of trials). Thus, we further examined whether a translation type effect was present when we excluded from analysis cases in which participants provided the same Arabic translation when they encountered the Hebrew word for the second time because repetition priming is likely to be a facilitator in such cases. The translation type effect in this analysis was again nonsignificant, $b = -0.16$, $SE = 0.56$, $z = -0.28$, $p = .78$, but the pattern of means was reversed

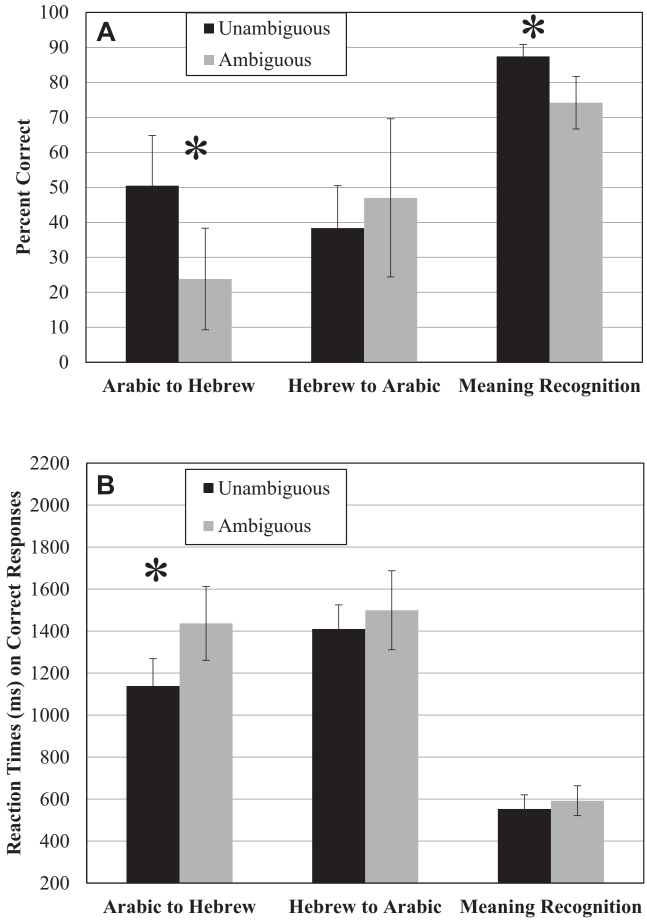

Figure 2 Response accuracy (A) and reaction time (B) as a function of translation type across the three tests. Error bars represent 95% confidence interval ($^{*}p < .05$).

($M_{\text{ambiguous}} = 0.37$, $M_{\text{unambiguous}} = 0.38$) such that participants responded to translation-ambiguous words less accurately than to translation-unambiguous words.

Individual Differences

Cognitive Resources

Phonological short-term memory contributed to learning such that an overall larger phonological short-term memory capacity was associated with marginally enhanced accuracy on the Hebrew-to-Arabic translation production test. More critically, phonological short-term memory modulated individuals' sensitivity to translation ambiguity during learning (see Figure 3).

For the Arabic-to-Hebrew translation production test, we observed a larger difference in both accuracy and RT measures between translation-ambiguous and translation-unambiguous items for individuals with larger phonological short-term memory compared to those with smaller phonological short-term memory (see Figure 3, Panel A and B). Whereas we observed the expected positive relation between phonological short-term memory and learning for translation-unambiguous items on the accuracy measure, for translation-ambiguous items, individuals with lower phonological short-term memory were as accurate as those with larger phonological short-term memory. We interpreted this finding to suggest that individuals with a greater phonological short-term memory capacity were unable to exhibit their learning advantage for translation-ambiguous items. This conclusion was further supported by the RT data. In particular, in the RT measure, individuals with larger phonological short-term memory took substantially longer to produce translation-ambiguous words than translation-unambiguous ones, whereas the difference for individuals with smaller phonological short-term memory was relatively small.

On the Hebrew-to-Arabic translation production test, there was no overall ambiguity disadvantage (see Figure 3, Panel C), and individuals with lower phonological short-term memory exhibited an advantage for ambiguous words. However, in this task as well, individuals with larger phonological short-term memory were able to capitalize on their resources for translation-unambiguous items more (i.e., steeper slope) than for translation-ambiguous ones, and thus were disadvantaged by translation ambiguity more than individuals with lower phonological short-term abilities. As was the case for the accuracy data of the Arabic-to-Hebrew translation production test, individuals with larger phonological short-term memory were unable to exhibit their learning advantage for translation-ambiguous items. This implied that translation ambiguity made learning more difficult, especially for participants with larger phonological short-term memory.

The variance of working memory that was not shared with phonological short-term memory (i.e., residualized working memory) did not modulate learning or the translation type effect. Similarly, residualized auditory statistical learning, as measured in our study, did not modulate learning. Finally, nonverbal intelligence, as measured by the Raven's Progressive Matrices, was positively linked to accuracy on the meaning recognition test, but critically could not serve as an alternative explanation for the observed effects of

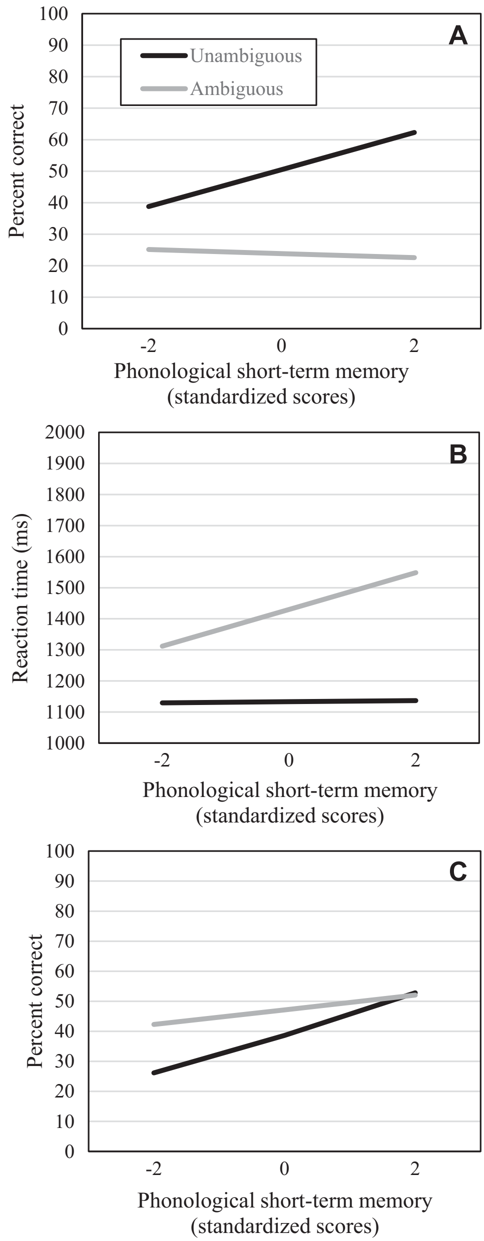

Figure 3 Translation type effect as a function of phonological short-term memory in the accuracy (A) and reaction time (B) of the Arabic-to-Hebrew translation production test and in the accuracy of the Hebrew-to-Arabic translation production test (C).

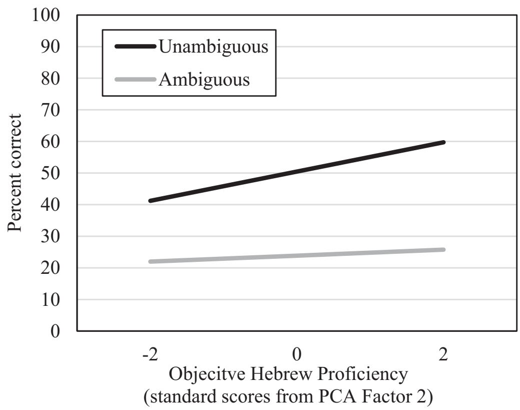

Figure 4 Translation type effect as a function of objective Hebrew proficiency in the accuracy data of the Arabic-to-Hebrew translation production test. PCA = principal component analysis.

translation type and phonological short-term memory because these effects were present even when variance related to nonverbal intelligence was partialled out by including it in the model.

Linguistic Background

Three orthogonal predictors were examined: the Multilingualism, Objective Hebrew, and Subjective Proficiency factors. Individual differences as measured by the Multilingualism factor did not modulate learning or the translation type effect in any of the examined measures. In contrast, proficiency in the language through which learning took place (i.e., Hebrew) was associated with improved performance (in accuracy and marginally in RT) in the meaning recognition test. Further, the Objective Hebrew Proficiency factor significantly interacted with the translation type effect for accuracy on the Arabic-to-Hebrew translation production test. As Figure 4 shows, the difference between translation-ambiguous and translation-unambiguous items was more pronounced with increased proficiency in Hebrew.

Finally, the effect of the Subjective Proficiency factor was positively linked to accuracy for the two translation production tests. In the principal component analysis described previously, this factor most strongly represented subjective ratings of proficiency in both Hebrew and additional languages and might reflect participants' bias in self-reporting or a shared proficiency construct, as suggested by the linguistic interdependence hypothesis (Cummins, 1979) and the common underlying cognitive processes proposal (Geva, 2014). We return to this issue in the discussion section.

Effect of Time and Response Type

To determine the contribution of the four-level variable of session in the Arabic-to-Hebrew translation production test, we adopted model comparisons using log likelihood ratio tests. These showed that the model including the main effect of session was superior to the one not including it, $\chi^2(3) = 94.96$, $p < .001$, but that adding the interaction between session and translation type did not improve the fit, $\chi^2(3) = 2.16$, $p = .54$. Thus, learning performance significantly differed across sessions for the Arabic-to-Hebrew translation production test. Follow-up tests with the Bonferroni corrections for multiple comparisons revealed that performance significantly differed across sessions (all *p*s $< .01$), with the exception of Sessions 3 and 4, which did not differ from each other, and Sessions 2 and 4, which were marginally different. In particular, performance improved between Session 1 ($M = 0.31$) and Session 2 ($M = 0.48$) but later deteriorated somewhat in Session 3 ($M = 0.42$) and Session 4 ($M = 0.43$). A similar pattern emerged in the RT data for this test, such that all sessions differed from each other (all *p*s $< .001$), with the exception of Sessions 1 and 3, as well as 2 and 4, which did not differ from each other. As for the accuracy data, performance improved between Session 1 ($M =$ 1,304 milliseconds) and Session 2 ($M =$ 1,121 milliseconds), and then deteriorated in Session 3 ($M =$ 1,271 milliseconds) and Session 4 ($M =$ 1,160 milliseconds). This pattern can be readily explained by the overall design of the study (see Figure 1 above), by which learning cycles were included in the first two sessions only. Critically, the translation type effect was stable across sessions.

In the Hebrew-to-Arabic translation production test, administered only in Sessions 3 and 4, there was a significant improvement between sessions in both accuracy ($M = 0.38$ vs. $M = 0.44$) and RT ($M =$ 1,482 milliseconds vs. $M =$ 1,400 milliseconds). This pattern was consistent with that found in previous research showing a significant increase in performance in a translation production task between sessions that did not include learning cycles (Degani et al., 2014) and may reflect consolidation processes (Batterink, Westerberg, & Paller, 2017). The translation type effect was not present and was not modulated by session.

For the meaning recognition test, we included response type (*no* vs. *yes* responses) as an additional fixed effect in the model to control for potential response bias due to the dichotomous nature of the task. Response type significantly modulated performance for the accuracy data, such that *yes* responses ($M = 0.78$) were made less accurately than *no* responses ($M = 0.87$). Critically, the translation type effect was not modulated by this variable.

Additional Post Hoc Analyses

Null Effect of Multilingualism

The lack of a multilingualism effect in our study was surprising given the documented advantage of multilingualism in novel word learning (Hirosh & Degani, 2018). However, Bogulski et al. (2018) found that such an advantage was not present when multilinguals were learning through their L2, although from a post hoc analysis, they suggested that, if that L2 had become the dominant language, an advantage might be present. In our study, all participants learned the foreign language (Arabic) through their dominant language (Hebrew), but for the native Russian speakers, this was their L2 rather than their L1. Thus, it is possible that the multilingualism effect would be limited to learners learning through their L1 (i.e., native Hebrew speakers) and would not be extended to any dominant language. To examine whether the multilingualism effect is dependent on the order of acquisition of the language from which learning takes place, we tested whether the multilingualism effect interacted with the L1 of our participants. We followed the same model structure, with the addition of L1 group (native Hebrew vs. native Russian) and the interaction of L1 group with the Multilingualism factor as fixed effects. For the Arabic-to-Hebrew translation production test, the effect of the Multilingualism factor and of the interaction between it and L1 group were not significant (ps $> .20$). For the Hebrew-to-Arabic translation production test as well as for the meaning recognition test, which were based on substantially fewer observations, the accuracy analyses failed to converge, and the effects were not significant in the RT analyses (ps $> .60$). Therefore, within our sample and as operationalized in our experiment, multilingualism did not modulate learning.

Discussion

In our study, we examined the separate and combined effects of word characteristics and learner characteristics on auditory foreign language vocabulary learning involving languages with different scripts. Specifically, we focused on translation ambiguity and on the way in which individual differences in cognitive resources and linguistic background (including language proficiency and multilingual language experience) modulate the difficulty of learning translation-ambiguous words. To this end, native Hebrew speakers and Russian–Hebrew speakers learned translation-unambiguous and translation-ambiguous Arabic words and were tested over four sessions in translation production and meaning recognition tests.

Translation Ambiguity Effect

In accordance with previous work on foreign language learning (Degani & Tokowicz, 2010; Degani et al., 2014), the results of our study demonstrated an ambiguity disadvantage in learning, so that backward translation production of translation-ambiguous words was slower and less accurate than that of translation-unambiguous words. We observed a similar disadvantage for the accuracy data of the meaning recognition test, but not of the forward translation production test. Further, the disadvantage did not dissipate with time (see also Degani & Tokowicz, 2010; Degani et al., 2014), reflecting the longer-term nature of the effect. Extending previous findings, we found a translation-ambiguity disadvantage in our study for foreign language words presented auditorily. Hence, this effect can be traced to the phonological level, and does not rely on competition between alternative orthographic representations or on difficulty in mapping competing phonologies onto the same orthography (see also Kaushanskaya & Marian, 2009). Our study thus provides clear evidence for a general difficulty in learning translation-ambiguous words.

Interestingly, the disadvantage was not present in a task that required retrieval of one of two newly learned phonological codes (Hebrew-to-Arabic translation production) but was strong in tasks that tapped the mapping of the foreign language to the known language (Arabic-to-Hebrew translation production) and that tapped the mapping of the foreign language to its corresponding meaning representation (meaning recognition). These findings suggest that the disadvantage associated with learning translation-ambiguous words is rooted in the many-to-one mapping of the foreign language phonology to the learner's existing representations (Bracken et al., 2017; Degani & Tokowicz, 2010) rather than in the competition between alternative phonological forms during retrieval. The lack of a reliable translation-ambiguity effect in the forward translation task coincides with patterns observed in previous studies (Degani et al., 2014; Degani & Tokowicz, 2010) and underscores the weak contribution of competition during retrieval as the underlying cause of the translation-ambiguity effect. However, the lack of translation-ambiguity effect in the forward translation task might also have been due to the overall poor performance on this task. This performance is likely to improve with increased foreign language proficiency. It is possible that, as performance in forward translation production improves, the competition between associated foreign language translations will begin to affect performance (see Prior, Kroll, & MacWhinney, 2013, for translation-ambiguity effects in more proficient bilinguals). Such competition processes will add to the difficulty created by the one-to-many mapping, such that both reduced associative strength and competition processes will hinder performance on translation-ambiguous words, leading to a more pronounced translation-ambiguity disadvantage.

Phonological Short-Term Memory

A second goal of our study was to test the extent to which the disadvantage associated with learning translation-ambiguous words is modulated by individual differences in cognitive resources and linguistic background. In particular, previous studies observed a correlation between adults' phonological short-term memory capacity and their ability to learn novel vocabulary (Martin & Ellis, 2012). In our study, enhanced phonological short-term memory was associated with better performance only on a task that directly required production of the novel foreign language words. In contrast, on tasks that tapped the mapping of the novel word to its Hebrew translation or to its meaning, we observed no modulation. The presence of phonological short-term memory modulations may therefore depend on the degree of difficulty posed by the task. First, as reflected by the pattern of means, the Hebrew-to-Arabic translation production test was overall a more demanding task. Further, even in the less demanding task (i.e., the Arabic-to-Hebrew translation production test), when task difficulty was enhanced due to stimuli characteristics (i.e., translation-ambiguous items), we observed phonological short-term memory modulations. We thus predict that individual differences in phonological short-term memory, and cognitive resources more generally, will exhibit larger effects when task demands become higher. One way in which foreign language vocabulary learning may become more difficult is when the to-be-learned foreign language words include nonnative phonemes because previous work has suggested that the advantage associated with increased phonological short-term memory is restricted to items composed of nonnative phonemes (Kaushanskaya, 2012, but see Martin & Ellis, 2012, for effects with native phonemes). In our study, the to-be-learned foreign language (Arabic) words included both novel and familiar phonemes to the relevant population of Hebrew speakers (see Saiegh-Haddad & Henkin-Roitfarb, 2014). It is possible that stronger effects of phonological short-term memory would be observed when the to-be-learned L2 is more phonologically dissimilar.

More critically, our study examined whether phonological short-term memory specifically modulated participants' ability to learn translation-ambiguous words. In particular, as noted above, Degani and Tokowicz (2010) suggested that translation-ambiguous words may be more difficult to learn due to active competition among the alternative foreign language translations or due to reduced associative strength similar to the fan effect. According to the competition account, when learners are asked to retrieve the newly learned foreign language word, the availability of the phonological form of the alternative foreign language translation may actively compete during the retrieval process, and such competition may be stronger for individuals who are able to retrieve both phonological forms more easily due to increased phonological short-term memory. Such an account would predict, however, that stronger competition would be observed in the forward translation direction when learners are required to select one of the two alternative foreign language translations.

Our study revealed no overall translation-ambiguity effect in the forward translation task and only minimal phonological short-term memory modulations in this translation direction. Alternatively, under the reduced associative strength explanation (i.e., fan effect), the one-to-many mapping leads to reduced associative strength between multiple linked representations. Individuals with increased phonological short-term memory may be better able to create and maintain multiple representations and, consequently, suffer a greater disadvantage than individuals who are less successful in establishing multiple linked associations. In this view, competition between phonological forms is not a critical component of the translation-ambiguity effect, and thus the effects are not limited to the forward translation direction. In accordance with this prediction, we observed that enhanced phonological short-term memory was associated with a larger translation-ambiguity disadvantage in the backward translation production test for both RT and accuracy measures. Specifically, we observed larger accuracy and RT gaps between translation-ambiguous and translation-unambiguous items for individuals with increased phonological short-term memory.

Based on previous studies (Kaushanskaya, 2012; Martin & Ellis, 2012), we expected individuals with larger phonological short-term memory to outperform those with smaller phonological short-term memory. As Figure 3 shows, these individuals indeed exhibited increased learning accuracy, but only in the case of translation-unambiguous items, whereas for translation-ambiguous items, there was no associated advantage for increased phonological short-term memory. This suggests that individuals with increased phonological short-term memory experienced translation-ambiguous items as more difficult to learn and were unable to capitalize on their resources to exhibit their full learning potential on such items. This conclusion is further supported by the latency data in which translation-ambiguous items were associated with longer RTs for individuals with larger phonological short-term memory compared to those with smaller phonological short-term memory. Together, these findings suggest that the translation-ambiguity disadvantage is more pronounced for those individuals who are better able to establish and maintain multiple linked representations (i.e., one-to-many mappings).

Additional Cognitive Resources

In our study, individual differences in working memory did not modulate learning of foreign language words. Critically, because of the overlap between phonological short-term memory and working memory at both the theoretical level (e.g., Baddeley, 2003; Buac et al., 2016) and the observed level, we partialled out the phonological storage component from the working-memory measure that we included such that it included only variance not explained by phonological short-term memory and related to the manipulation of information. This more restricted measure did not modulate learning or the translation-ambiguity disadvantage. It is possible, however, that a more heterogeneous group of learners in terms of cognitive resources, or a more demanding learning protocol with greater task demands would give rise to learning modulations by working memory.

In addition, individual differences in auditory statistical learning did not contribute to the auditory foreign language learning in our study. Our version of the statistical learning task used verbal syllables of a novel language. Although we presented these stimuli as a novel language, it is possible that their processing relied to some extent on participants' prior linguistic knowledge (see Siegelman et al., 2017; Siegelman, Bogaerts, Elazar, Arciuli, & Frost, 2018). Thus, it is possible that a more recent version of the statistical learning task that is less reliant on participants' prior knowledge (e.g., utilizing nonverbal stimuli) would be better suited to elucidate the link between statistical learning abilities and adult foreign language auditory vocabulary learning.

Linguistic Background

In our study, prior knowledge was reflected in participants' linguistic background, including level of language proficiency in the dominant language and experience with additional languages. Two theoretical accounts have been proposed to explain observed links between prior linguistic abilities and learners' ability to learn a foreign language (e.g., Prior et al., 2014). According to the linguistic interdependence hypothesis, this link stems from the fact that learning both a L1 and additional languages rely on a common underlying proficiency/aptitude construct (Cummins, 1979). This common construct may be conceived of as including basic cognitive processes (Geva, 2014). Of relevance, according to this theoretical perspective, individuals with higher linguistic abilities in *any* existing languages are expected to outperform those with lower abilities. In contrast, according to the alternative theoretical approach, the typological/contrastive approach (Odlin, 1989), learners rely on their previous linguistic knowledge to the extent that it is similar to the to-be-learned language (see also MacWhinney, 2005). Positive transfer is therefore expected only between structurally similar languages.

The results of our study suggest that learners' proficiency in the more typologically similar language, Hebrew, was positively linked to foreign language learning, especially for the meaning recognition test. This is evident in the positive effect of the Objective Hebrew factor on the meaning recognition accuracy and RT performance. In contrast, learners' proficiency in languages other than Hebrew, as captured by the Multilingualism factor that we derived, was not correlated with learning in any of the measures. This pattern of results provides support for the typological/contrastive approach by which learners benefit from prior knowledge only when that prior knowledge resembles the to-be-learned language.

We should note several important caveats to this conclusion, however. First, whereas Hebrew proficiency was objectively measured using fluency tasks, proficiency in other languages was only subjectively estimated. Moreover, there was at least some shared variance in learners' subjective estimations of their proficiency in Hebrew and in other languages, as reflected by the Subjective Proficiency factor, and this shared variance influenced performance to some extent (i.e., the Hebrew-to-Arabic translation production test). Thus, to the extent that the Subjective Proficiency factor captured a shared proficiency construct, one may take the current findings as support for the interdependence hypothesis. Further, the association between basic cognitive resources such as phonological short-term memory and foreign language learning observed in our study may be taken to reflect an association between the common proficiency construct (Geva, 2014) and foreign language learning. In addition, and more importantly in our study, Hebrew was not only the more similar language, it was also the language through which learning took place. Thus, our study did not allow us to discern whether the facilitatory effect of Hebrew proficiency stemmed from its typological similarity to the foreign language, or from its status as the language through which learning took place. Future studies should address this more directly by manipulating the language through which foreign language learning occurs.

Interestingly, objective proficiency in Hebrew modulated the translation-ambiguity effect, such that individuals with higher Hebrew proficiency experienced a larger translation-ambiguity disadvantage. This pattern may be explained by the one-to-many mapping required for translation-ambiguous words. For individuals to learn such mapping, they need to reorganize their form-to-meaning space so that an ambiguous mapping is created. For example, in order to learn a translation-ambiguous item, a learner needs to establish two links from the existing meaning/translation in Hebrew (e.g., /makor/) to the two Arabic translations (/munqa:r/ and /masˤdar/). Notably, the more entrenched the representations in Hebrew are, the more difficult it is for a learner to rearrange these connections. As a result, individuals with higher proficiency in Hebrew are required to make more substantial changes in order to learn translation-ambiguous words.

As noted above, learners' proficiency in languages other than Hebrew, as captured by the Multilingualism factor, was not correlated with learning in any of the measures. The lack of multilingualism effect goes against the expectation based on previous studies (e.g., Bartolotti & Marian, 2012; Kaushanskaya & Rechtzigel, 2012; Kaushanskaya, Yoo, & Van Hecke, 2013; for review, see Hirosh & Degani, 2018; Tokowicz & Degani, 2015). Several explanations may account for this inconsistency. First, the majority of previous work examined the multilingualism effect when learning took place through participants' L1 (for review, see Hirosh & Degani, 2018). Recent work by Bogulski et al. (2018) suggested that the multilingual advantage in word learning is not expected when multilinguals learn through their L2. The authors suggested that multilinguals are better learners because of enhanced experience with regulating nontarget language activation and specifically with regulating their L1 in order to acquire new information. They further 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. In accordance with this refinement, all participants in our study learned the Arabic foreign language vocabulary through their dominant language (Hebrew) although this was the L2 for the Russian–Hebrew speakers. Nonetheless, we did not observe a multilingual advantage in foreign language learning. It is possible that in contrast to the suggestion of Bogulski et al. (2018) regarding language dominance, the order of acquisition of the languages does play an important role and that the multilingual advantage is limited to cases where learning is done through the first-acquired language irrespective of its dominance. In our sample, the order of acquisition did not modulate the (lack) of multilingualism effect, but larger samples may be more suited for directly testing this prediction.

A second explanation relates to the range of the Multilingualism factor sampled in our study. Whereas most previous research has compared bilingual speakers to monolingual speakers, our study examined the effect of degree of multilingualism in a heterogeneous group of native Hebrew speakers and native Russian speakers who had learned Hebrew. Critically, all participants were also proficient in English such that, although the Hebrew speakers were much less multilingual than the Russian–Hebrew speakers, all participants had some experience with foreign language learning. As a result, our sample did not include individuals with minimal or close to minimal experience with additional languages. This restricted range may have reduced our ability to observe a multilingualism effect, but notably may represent the typical distribution of language backgrounds in many countries (see also Kreiner & Degani, 2015). These different explanations are not mutually exclusive by any means. For instance, because all participants had some experience in learning English through Hebrew, they had gained experience in inhibiting the source language (Hebrew) in order to learn an additional language. Consequently, in our study, when we asked participants to learn Arabic using Hebrew as the source language, native Hebrew speakers and Russian–Hebrew speakers did not differ in their ability to do so.

Our study thus shows that multilinguals are not always better at foreign language learning and that the multilingual advantage may be constrained by the language in which learning takes place. In their review of the literature, Hirosh and Degani (2018) suggested that the multilingual advantage in novel language learning may stem from both direct and indirect sources. Specifically, direct sources of the multilingual advantage are dependent on the overlap in the specific previous experiences of the learner and the learning task. Multilinguals outperform bilinguals or monolinguals if they have had more experience in the specific learning situation, in this case regulating the activation of the source language (Hebrew) while learning a foreign language. According to this conceptualization, it is possible that multilingualism did not exert an effect in our study because all participants had experience in specifically regulating the source language (Hebrew) during foreign language learning (as Bogulski et al., 2018, have suggested is important). Participants further had knowledge of a typologically similar language (Hebrew) available to serve as a source of direct transfer. In addition to the direct effects of the type described above, Hirosh and Degani (2018) suggested that multilinguals may enjoy an advantage in foreign language learning mediated via indirect sources such as attentional control or sensitivity to statistical regularities. In our study, individual differences in cognitive resources were included as control variables in the model such that a multilingual effect could only reflect direct overlap between prior representations and learning experience and the foreign language vocabulary and learning situation. Nonetheless, it is possible that individual differences in other cognitive variables not examined here, such as attentional control (Yoshida, Tran, Benitez, & Kuwabara, 2011), may contribute indirectly to the multilingual advantage in foreign language learning. More research is needed to uncover the contribution of these and additional variables to explain how multilingualism affects novel language learning.

Limitations and Future Research

Three additional aspects of our study remain to be examined further. First, the translation-ambiguous words taught in our study were two Arabic words corresponding to a shared ambiguous Hebrew word. As a result, the two translations were largely unrelated in meaning, referred to as meaning-ambiguous words (Degani & Tokowicz, 2010; Eddington & Tokowicz, 2013) as opposed to form-ambiguous words, where the two foreign language translations are synonymous words (e.g., the Hebrew /sapa/ translates as both "couch" and "sofa"). However, recent work has suggested that a continuous measure of the meaning relatedness of the translations, termed translation semantic variability (Bracken et al., 2017) can capture substantial variability during foreign language learning. Future studies will thus examine whether meaning similarity of the translations affects auditory foreign language learning and to what extent this item characteristic interacts with learner individual differences.

Second, in our protocol, participants became aware during testing of the presence of translation-ambiguous words in the learning set. Further, ambiguous items were blocked and interleaved between two blocks of unambiguous words. Thus, it remains to be examined whether randomly interspacing ambiguous and unambiguous words within a single learning and testing block, a design that is likely to reduce the prominence of translation ambiguity, would change the way individuals approach learning. Moreover, although not analyzed here, we also exposed participants to ambiguous foreign language (Arabic) words during learning. Stimuli composition may have influenced learning because previous work has shown that the inclusion of translation-ambiguous words affects the processing of translation-unambiguous words (Tokowicz & Kroll, 2007).

Finally, overall accuracy in our study was relatively low, possibly due to the low number of repetitions of each word during learning. Specifically, similar previous research used 12 to 16 repetitions of each item (Degani et al., 2014; Degani & Tokowicz, 2010), whereas in our study each item was repeated only three times across the entire learning protocol. Accordingly, accuracy levels dropped relative to previous studies, but effective learning still took place, likely because additional changes were implemented in the protocol. Most critically, whereas in previous studies mere repetition was required during learning, in our study, participants engaged in active retrieval attempts, which have been shown to strengthen learning (Kang et al., 2013). As a result, accuracy levels two weeks after learning ranged from 33% to 47% in our study relative to 48–53% in Degani and Tokowicz (2010) and to 50–63% in Degani et al. (2014). Critically, in all three studies a robust translation-ambiguity disadvantage was observed, which was not alleviated by time. Nonetheless, all three studies tapped the beginning stages of learning, and the individual difference modulations observed in our study may not necessarily hold with increased proficiency in the foreign language. This possibility awaits future investigation at more advanced stages of foreign language learning.

Conclusion

To summarize, our study demonstrated the contribution of both item-based variables and learner-based variables to foreign language vocabulary learning. With respect to item characteristics, our study showed a robust and longer-term translation-ambiguity disadvantage during auditory foreign language vocabulary learning. The effect was mediated by phonology because the foreign language words were presented auditorily, and there was no script overlap between the foreign language and learners' known languages. As in previous studies, this effect did not diminish with time and was evident in both a translation production and a meaning recognition test. The presence of the effect across tests underscores its source in the one-to-many mappings required for translation-ambiguous words rather than active response competition.

With respect to learner characteristics, our study highlighted the significance of phonological short-term memory and proficiency in the language in which learning takes place. Rather than dichotomizing groups, our study illustrates the use of a more continuous measure of multilingual experience that is suited for use with heterogeneous and ecological samples of the population. In our sample, multilingual experience did not modulate learning. Interestingly, the study exemplifies interactions between item-based and learner-based characteristics such that individuals with higher phonological short-term memory and those with higher proficiency in the source language exhibited larger effects of translation-ambiguity during learning. These findings highlight the importance of examining both learner-based and item-based sources of variance in the same study and testing for the interactive nature of their influence.

Final revised version accepted 30 January 2019

References

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, *6*, 451–474. https://doi.org/10.1016/0010-0285(74)90021-8

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*, 417–423. https://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, *4*, 829–839. https://doi.org/10.1038/nrn1201

Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, *105*, 158–173. https://doi.org/10.1037/0033-295X.105.1.158

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

Batterink, L. J., Westerberg, C. E., & Paller, K. A. (2017). Vocabulary learning benefits from REM after slow-wave sleep. *Neurobiology of Learning and Memory*, *144*, 102–113. https://doi.org/10.1016/j.nlm.2017.07.001

Bialystok, E., McBride-Chang, C., & Luk, G. (2005). Bilingualism, language proficiency, and learning to read in two writing systems. *Journal of Educational Psychology*, *97*, 580–590. https://doi.org/10.1037/0022-0663.97.4.580

Boada, R., Sánchez-Casas, R., Gavilán, J. M., García-Albea, J. E., & Tokowicz, N. (2012). Effect of multiple translations and cognate status on translation recognition performance of balanced bilinguals. *Bilingualism: Language and Cognition*, *16*, 183–197. https://doi.org/10.1017/S1366728912000223

Bogulski, C. A., Bice, K., & Kroll, J. F. A. (2018). Bilingualism as a desirable difficulty: Advantages in word learning depend on regulation of the dominant language. *Bilingualism: Language and Cognition*. Published online 10 August 2018. https://doi.org/10.1017/S1366728918000858

Bracken, J., Degani, T., Eddington, C., & Tokowicz, N. (2017). Translation semantic variability: How semantic relatedness affects learning of translation ambiguous words. *Bilingualism: Language and Cognition*, *20*, 783–794. https://doi.org/10.1017/S1366728916000274

Buac, M., Gross, M., & Kaushanskaya, M. (2016). Predictors of processing-based task performance in bilingual and monolingual children. *Journal of Communication Disorders*, *62*, 12–29. https://doi.org/10.1016/j.jcomdis.2016.04.001

Cenoz, J., Hufeisen, B., & Jessner, U. (2003). Why investigate the multilingual lexicon? In J. Cenuz, B. Hufeisen, & U. Jessner (Eds.), *The multilingual lexicon* (pp. 1–9). Dordrecht, Netherlands: Kluwer. https://doi.org/10.1007/978-0-306-48367-7_1

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

Cummins, J. (1991). Interdependence of first-and second-language proficiency in bilingual children. *Review of Educational Research*, *9*, 222–251. https://doi.org/10.1017/CBO9780511620652.006

de Groot, A. M. B., & van Hell, J. G. (2005). The learning of foreign language vocabulary. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 9–29). New York, NY: Oxford University Press.

Degani, T. (2011). Bidirectional transfer: Consequences of translation ambiguity for bilingual word meaning (Unpublished doctoral dissertation). University of Pittsburgh, Pittsburgh, PA.

Degani, T., Prior, A., Eddington, C. M., da Luz Fontes, A. B. A., & Tokowicz, N. (2016). Determinants of translation ambiguity. *Linguistic Approaches to Bilingualism*, *6*, 290–307. https://doi.org/10.1075/lab.14013.deg

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

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

Degani, T., & Tokowicz, N. (2013). Cross-language influences: Translation status affects intra-word sense relatedness. *Memory & Cognition*, *41*, 1046–1064. https://doi.org/10.3758/s13421-013-0322-9

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

Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (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

Doherty, M. J. (2004). Children's difficulty in learning homonyms. *Journal of Child Language*, *31*, 203–214. https://doi.org/10.1017/S030500090300583X

Eddington, C. M., & Tokowicz, N. (2013). Examining English–German translation ambiguity using primed translation recognition. *Bilingualism: Language and Cognition*, *16*, 442–457. https://doi.org/10.1017/S1366728912000387

Edwards, J. V. (2004). Foundations of bilingualism. In T. K. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism* (pp. 7–31). Malden, MA: Blackwell.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331. https://doi.org/10.1037/0096-3445.128.3.309

Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review*, *37*, 66–108. https://doi.org/10.1016/j.dr.2015.05.002

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

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences*, *19*, 117–125. https://doi.org/10.1016/j.tics.2014.12.010

Frost, R., Siegelman, N., Narkiss, A., & Afek, L. (2013). What predicts successful literacy acquisition in a second language? *Psychological Science*, *24*, 1243–1252. https://doi.org/10.1177/0956797612472207

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education*, *8*, 259–272. https://doi.org/10.1007/BF03174081

Gathercole, S. E., Pickering, S. J., Hall, M., & Peaker, S. M. (2001). Dissociable lexical and phonological influences on serial recognition and serial recall. *Quarterly Journal of Experimental Psychology*, *54*, 1–30. https://doi.org/10.1080/02724980042000002

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

Goral, M. (2018). The bilingual mental lexicon beyond Dutch–English written words. *Bilingualism: Language and Cognition*, 1–2. Published online 18 June 2018. https://doi.org/10.1017/S1366728918000743

Harrington, M., & Sawyer, M. (1992). L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, *14*, 25–38. https://doi.org/10.1017/S0272263100010457

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

Hummel, K. M., & French, L. M. (2010). Phonological memory and implications for the second-language classroom. *The Canadian Modern Language Review*, *66*, 371–391. https://doi.org/10.3138/cmlr.66.3.371

Jiang, N. (2018). Phonology-based bilingual activation among different-script bilinguals? *Bilingualism: Language and Cognition*, 1–2. Published online 19 June 2018. https://doi.org/10.1017/S1366728918000664

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

Karpicke, J. D., & Roediger, H. L. III. (2008). The critical importance of retrieval for learning. *Science*, *319*, 966–968. https://doi.org/10.1126/science.1152408

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

Kaushanskaya, M., Gross, M., & Buac, M. (2014). Effects of classroom bilingualism on task-shifting, verbal memory, and word learning in children. *Developmental Science*, *17*, 564–583. https://doi.org/10.1111/desc.12142

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

Kaushanskaya, M., Marian, V., & Yoo, J. (2011). Gender differences in adult word learning. *Acta Psychologica*, *137*, 24–35. https://doi.org/10.1016/j.actpsy.2011.02.002

Kaushanskaya, M., & Rechtzigel, K. (2012). Concreteness effects in bilingual and monolingual word learning. *Psychonomic Bulletin & Review*, *19*, 935–941. https://doi.org/10.3758/s13423-012-0271-5

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

Kaushanskaya, M., Yoo, J., & Van Hecke, S. (2013). Word learning in adults with second-language experience: Effects of phonological and referent familiarity. *Journal of Speech, Language, and Hearing Research*, *56*, 667–678. https://doi.org/10.1044/1092-4388(2012/11-0084)

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

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

Kilgarriff, A., Reddy, S., Pomikálek, J., & Avinesh, P. V. S. (2010). A corpus factory for many languages. In N. Calzolari, K. Choukri, B. Maegaard, J. Odijk, S. Piperidia, M. Rosner, & D. Tapias (Eds.), *Proceedings of the Seventh International Conference on Language Resources and Evaluation* (pp. 904–910). Valetta, Malta: European Language Resources Association.

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

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

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*, 31–69. https://doi.org/10.1111/1467-9922.00032

Lu, Y., Wu, J., Dunlap, S., & Chen, B. (2017). The inhibitory mechanism in learning ambiguous words in a second language. *Frontiers in Psychology*, *8.636*. https://doi.org/10.3389/fpsyg.2017.00636

MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). New York, NY: Oxford University Press.

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*, 940–967. https://doi.org/10.1044/1092-4388(2007/067)

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

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception. Part 1: An account of basic findings. *Psychological Review*, *88*, 375–407. https://doi.org/10.1037/0033-295X.88.5.375

Melby-Lervåg, M., & Lervåg, A. (2011). Cross-linguistic transfer of oral language, decoding, phonological awareness and reading comprehension: A meta-analysis of the correlational evidence. *Journal of Research in Reading*, *34*, 114–135. https://doi.org/10.1111/j.1467-9817.2010.01477.x

Michael, E. B., Tokowicz, N., Degani, T., & Smith, C. J. (2011). Individual differences in the ability to resolve translation ambiguity across languages. *Vigo International Journal of Applied Linguistics*, *8*, 79–97.

Miyake, A., Just, M. A., & Carpenter, P. A. (1994). Working memory constraints on the resolution of lexical ambiguity: Maintaining multiple interpretations in neutral contexts. *Journal of Memory and Language*, *33*, 175–202. https://doi.org/10.1006/jmla.1994.1009

Norman, T., Degani, T., & Peleg, O. (2016). Transfer of L1 visual word recognition strategies during early stages of L2 learning: Evidence from Hebrew learners whose first language is either Semitic or Indo-European. *Second Language Research*, *32*, 109–122. https://doi.org/10.1177/0267658315608913

Odlin, T. (1989). *Language transfer: Cross-linguistic influence in language learning*. Cambridge, UK: Cambridge University Press.

Pasquarella, A., Gottardo, A., Grant, A. (2012). Comparing factors related to reading comprehension in adolescents who speak English as a first (L1) or second (L2) language. *Scientific Studies of Reading*, *16*, 475–503. https://doi.org/10.1080/10888438.2011.593066

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

Prior, A., Goldina, A., Shany, M., Geva, E., & Katzir, T. (2014). Lexical inference in L2: Predictive roles of vocabulary knowledge and reading skill beyond reading comprehension. *Reading and Writing*, *27*, 1467–1484. https://doi.org/10.1007/s11145-014-9501-8

Prior, A., Kroll, J. F., & MacWhinney, B. (2013). Translation ambiguity but not word class predicts translation performance. *Bilingualism: Language and Cognition*, *16*, 458–474. https://doi.org/10.1017/S1366728912000272

R Core Team. (2017). R: A language and environment for statistical computing (Version 3.3.3) [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org

Raven, J. C. (1960). *Guide to the standard progressive matrices: Sets A, B, C, D and E*. London, UK: HK Lewis.

Raven, J. C., & Rust, J. (2008). *Standard Progressive Matrices-Plus Version and Mill Hill Vocabulary Scale: Manual*. London, UK: Pearson.

Raviv, L., & Arnon, I. (2018). The developmental trajectory of children's auditory and visual statistical learning abilities: Modality-based differences in the effect of age. *Developmental Science*, *21*, 1–13. https://doi.org/10.1111/desc.12593

Rodd, J. M., Berriman, R., Landau, M., Lee, T., Ho, C., Gaskell, G. M., & 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

Saiegh-Haddad, E., & Geva, E. (2008). Morphological awareness, phonological awareness, and reading in English-Arabic bilingual children. *Reading and Writing*, *21*, 481. https://doi.org/10.1007/s11145-007-9074-x

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

Share, D. L. (2008). On the Anglocentricities of current reading research and practice: The perils of overreliance on an "outlier" orthography. *Psychological Bulletin*, *134*, 584–615. https://doi.org/10.1037/0033-2909.134.4.584

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

Shimron, J. (2003). Semitic languages: Are they really root-based? In J. Shimron (Ed.), *Language processing and acquisition in languages of Semitic, root-based morphology* (pp. 1–28). Amsterdam, Netherlands: John Benjamins.

Siegelman, N., Bogaerts, L., Christiansen, M. H., & Frost, R. (2017). Towards a theory of individual differences in statistical learning. *Philosophical Transactions of the Royal Society B*, *372*, 1–10. https://doi.org/10.1098/rstb.2016.0059

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition*, *177*, 198–213. https://doi.org/10.1016/j.cognition.2018.04.011

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language*, *81*, 105–120. https://doi.org/10.1016/j.jml.2015.02.001

Sparks, R. L., Patton, J., Ganschow, L., & Humbach, N. (2012). Relationships among L1 print exposure and early L1 literacy skills, L2 aptitude, and L2 proficiency. *Reading and Writing*, *25*, 1599–1634. https://doi.org/10.1007/s11145-011-9335-6

Tokowicz, N. (2014). Translation ambiguity affects language processing, learning, and representation. In R. T. Miller, K. I. Martin, C. M. Eddington, A. Henery, N. Marcos Miguel, A. M. Tseng, A. Tuninetti, & D. Walter (Eds.), *Selected proceedings of the 2012 Second Language Research Forum: Building bridges between disciplines* (pp. 170–180). Somerville, MA: Cascadilla Press.

Tokowicz, N., & Degani, T. (2010). Translation ambiguity: Consequences for learning and processing. In B. VanPatten & J. Jegerski (Eds.), *Research on second language processing and parsing* (pp. 281–293). Amsterdam, Netherlands: John Benjamins.

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.

Tokowicz, N., & Kroll, J. F. (2007). Number of meanings and concreteness: Consequences of ambiguity within and across languages. *Language and Cognitive Processes*, *22*, 727–779. https://doi.org/10.1080/01690960601057068

Tokowicz, N., Kroll, J. F., de Groot, A. M. B., & van Hell, J. G. (2002). Number-of-translation norms for Dutch-English translation pairs: A new tool for examining language production. *Behavior Research Methods, Instruments, & Computers*, *34*, 435–451. https://doi.org/10.3758/BF03195472

Tomoschuk, B., Ferreira, V. S., & Gollan, T. H. (2018). When a seven is not a seven: Self-ratings of bilingual language proficiency differ between and within language populations. *Bilingualism: Language and Cognition*. Published online 13 June 2018. https://doi.org/10.1017/S1366728918000421

van Heuven, W. J., & Wen, Y. (2018). The need for a universal computational model of bilingual word recognition and word translation. *Bilingualism: Language and Cognition*, 1–2. Published online 27 June 2018. https://doi.org/10.1017/S1366728918000688

Wechsler, D. (1997). *Wechsler Memory Scale (WMS–III)* (Vol. 14). San Antonio, TX: Psychological Corporation.

Wen, Z. E., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, *50*, 1–31. https://doi.org/10.1017/S0261444816000276

Yoshida, H., Tran, D. N., Benitez, V., & Kuwabara, M. (2011). Inhibition and adjective learning in bilingual and monolingual children. *Frontiers in Psychology*, *2*, 210. https://doi.org/10.3389/fpsyg.2011.00210

Supporting Information

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

Appendix S1. List of Stimuli and Definitions by Translation Type.
Appendix S2. Mixed-Effects Model Estimates Across Measures.

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

How Learners Cope With Ambiguity in Foreign Language Vocabulary Learning

What This Research Was About And Why It Is Important

Words often have multiple or vague meanings that do not translate easily to other languages. As a result, a single word in one language may translate into two correct translations in another language. For instance, the Hebrew word /makor/ means both "beak" and "source." Previous research has shown that such translation ambiguity (having to learn separate words that map onto a single word with multiple meanings in an already known language) creates difficulty during vocabulary learning. The researchers extended this line of research to test whether such difficulty is observed when the to-be-learned foreign language words are spoken rather than presented visually (written out). The researchers further examined whether learning of foreign language vocabulary is affected by learners' cognitive resources (e.g., their memory capacity) and linguistic background (e.g., their language proficiency and multilingual language experience).

What the Researchers Did

- The researchers tested 53 native Hebrew speakers and Russian–Hebrew speakers (all of whom also knew English). The participants learned previously unknown Arabic words; these words were presented auditorily, and the participants saw the definition of these words (in Hebrew) along with their translation (also in Hebrew).

- In some cases, two Arabic words corresponded to the same ambiguous Hebrew translation (e.g., /munqa:r/ meaning "beak" and /masˤdar/ meaning "source" were learned in association with the ambiguous Hebrew word /makor/). Such translation-ambiguous words were compared to unambiguous items, where one meaning in Hebrew corresponded to one meaning in Arabic.
- During learning, which took place over two days, the participants repeated the Arabic words and also recalled them in response to the Hebrew translation, before hearing the correct Arabic word.
- During testing, the participants produced the Hebrew translation after hearing the Arabic word and produced the Arabic word after seeing the Hebrew translation. The participants further decided if a Hebrew definition corresponded to the Arabic word they heard.
- The participants also performed additional tasks measuring their cognitive resources (e.g., memory capacity) and their language background.

What the Researchers Found

- Translation-ambiguous words were more difficult to learn than translation-unambiguous words, which was the case both immediately after learning and one and three weeks after learning.
- The difficulty of learning translation-ambiguous words did not dissipate with time. This difficulty appears to be linked to how participants store and access the sounds of words, not their spelling.
- The participants with a larger memory capacity for sounds demonstrated better learning than those with a smaller memory capacity, but not in the case of translation-ambiguous words.
- The participants' proficiency in Hebrew (i.e., the language they used for learning new words), but not their multilingualism (i.e., also speaking Russian or English besides Hebrew), mattered for vocabulary learning.

Things to Consider

- Having to learn two translations for the same word might create a challenge that is equally difficult for learners with more and fewer cognitive resources.
- Learners might capitalize on their prior linguistic knowledge for vocabulary learning, but multilinguals are not always better in learning vocabulary in additional languages.
- Most importantly, not all words are equally easy to learn, and not all learners approach vocabulary learning with similar capacities.

How to cite this summary: Degani, T., & Goldberg, M. (2019). How learners cope with ambiguity in foreign language vocabulary learning *OASIS Summary* of Degani & Goldberg in *Language Learning*. https://oasis-database.org