

(Not) acquiring meaning in a second language: Are input deficits key?

Barbara C. Malt, Xingjian Yang, and Jessica Joseph^{1*}

Abstract. Word meanings are not always parallel across languages, and second language (L2) learners often use words in non-native ways. Is the learning problem inherent in maintaining conflicting word-to-meaning mappings within an integrated lexical network, or is it due to insufficient attention to and input for acquiring L2 mappings? To help discriminate between these possibilities, we gave English speakers repeated exposures to 40 brief videos of actions, labeled with five novel words that cross-cut English labeling patterns. Half the participants were told only to learn the labels for the actions. The other half were told to figure out their meanings, which might differ from English. The Figure Out Meanings group made test choices faster and were also slightly more likely to produce definitions capturing the intended meanings. However, both groups performed well above chance in generalizing the novel words. High levels of choice performance for both groups point to insufficient input, rather than inherent properties of lexical networks, as the critical limiting factor in more typical L2 learning contexts. Speed and definition performance hint at some advantage to explicit attention in sorting out L1-L2 differences.

Keywords. word learning; word meaning; vocabulary; second language learning; L1 influence on L2

1. Introduction. Languages are diverse in their lexicons as well as their grammar and sound systems. For instance, English labels a ball on a table and a handle on a door both as *on*, whereas Dutch labels them separately as *op* vs. *aan*. On the other hand, Spanish labels both along with an apple in a bowl as *en* (Bowerman 1996). Languages also vary in how many color terms they use to divide up the spectrum, what distinctions are drawn among drinking vessels, and the meaning of body part terms, among others (for review, see Malt & Majid 2013). These differences are not easily mastered by second language (L2) learners. They often use words in non-native ways even after many years of L2 immersion (De Groot 2012, Malt & Sloman 2003, Pavlenko 2009).

The challenge for the L2 learner entails acquiring lexical categories that may only partially overlap with first language (L1) categories, may be subsets or supersets of L1 categories, or may cross-cut them by using entirely different semantic dimensions. In lab studies of artificial category learning, learning one set of categories followed by learning a cross-cutting set produces large costs in speed and accuracy of categorization and perseverative errors (e.g., Kruschke 1996). In L2 acquisition in natural (non-laboratory) contexts, cross-cutting also proves difficult. For instance, English speakers *carry* or *hold* an object, regardless of manner of support or contact, whereas in Korean, Chinese, and Japanese, the same actions are labeled according to manner of contact or support, regardless of movement (Saji, Imai, Saalbach, Zhang, Shu & Okada 2011). Malt (2020) studied native Mandarin speakers' use of English *carry* and *hold* and found that even Mandarin speakers immersed in an English language environment showed persistent errors in the use of these terms (see also Jessen & Cadierno 2013, Malt, Li, Pavlenko, Zhu & Ameer 2015).

^{1*} We thank Ashley Veliz for assistance with data collection and coding and Helen Borchart and YiChen Shen for serving video models. Authors: Barbara Malt, Lehigh University (bcm0@Lehigh.edu), Xingjian Yang, Lehigh University (xij221@Lehigh.edu), Jessica Joseph, Lehigh University (jlj217@lehigh.edu)

Such results suggest that entrenchment in an initial set of categories interferes with encoding alternative groupings of the same entities (for discussion of entrenchment in language learning, see, e.g., MacWhinney 2017, Schmid 2017). In lexical network terms, once the network settles into a stable configuration of links between word forms and elements of meaning, remapping may become difficult. Furthermore, continued use of the L1 may create a continued pull away from the L2 form-to-meaning mappings (Malt et al., 2015).

Paradoxically, though, within the L1, language users easily master words having complex relations among their meanings. For instance, *cup* both overlaps with *mug* and can be used to encompass mugs, and goal-derived and thematic terms cross-cut taxonomic ones (e.g., *pet* and *wildlife* crosscut *feline* and *canine*; similarly for *breakfast foods* and *dinner foods* vs. *dairy* and *grains* or *things to take on a picnic* vs. *drinks* and *utensils* (Barsalou 1991, Ross & Murphy 1999). L1 learners acquire and maintain many such cross-cutting terms without difficulty. These observations indicate that at least within a single language, there is not a fundamental difficulty in simultaneously representing different ways of grouping the same entities by name.

The contrast between ease of acquiring cross-cutting mappings in L1 and persistent difficulties in L2 raises the question of where the difference lies. Is the L2 problem somehow inherent in how the lexical network deals with conflicting L1 and L2 mappings? Or could it be due to differences in how L1 and L2 are acquired? Young L1 learners often are given many exposures to examples of the same word-referent pairs, and the input often encourages them to focus on extracting meaning (e.g., Look, I have a *lion*. Look at the *lion*. Do you know what a *lion* is?; Snow, 1972). Both of these input features may facilitate extraction of meaning. L2 learners, especially outside of a classroom, are likely to hear much more diverse input and to need to focus on understanding the larger discourse goals. Child L1 learners also have metalinguistic knowledge that meanings are contrastive (Clark, 1987; that is, that *kitty* and *lion* are different, and that *pet* and *wildlife* should differ from *feline* and *canine*), whereas L2 learners may simply assume that L2 meanings are parallel to L1 meanings. This is likely true in both formal instruction and in immersion contexts.

The current study was designed to discriminate between the possibility that learning difficulty is due to fundamental characteristics of lexical networks and the possibility that it is due to insufficient input and attention to the input. It also investigated whether metalinguistic knowledge that L2 meanings can differ is crucial to successful learning.

We used a domain where word meanings cross-cut each other in the L1 and the to-be-learned language. We gave participants intensive L2 word-referent pairing exposure and tested their ability to define the L2 words, label exposed instances, and generalize to new instances. Further, half the participants were told simply to learn the L2 labels and half were told that meanings may differ from L1, thereby cueing them to search for new meanings.

2. Method.

2.1. PARTICIPANTS. Seventy-eight native English speakers without substantial knowledge of another language participated. They were recruited from Lehigh University's introductory psychology participant pool. Native language and other language experience was assessed with a brief language history questionnaire. Data from an additional five participants with exposure to Asian languages, Arabic, and Urdu were not included.

2.2. STIMULI. Stimuli were designed to have meanings cross-cutting the meanings of English words for them. Training stimuli were videos illustrating five novel verbs for actions of standing or walking with an object, modeled after Saji et al. (2011). The actions are labeled as *carry* or *hold* in English, as established in a pre-test. In many Asian languages and in our lab version, these

actions are labeled according to the manner of the object being in contact with or supported by the person (e.g., cupped in one or both palms vs. held snug against the front or side of the body), regardless of whether the action is stationary or moving (see Saji et al. 2011).

We created 40 training videos to illustrate 5 novel verbs of interacting with objects. The verbs were based on Korean verbs for this domain (as described in Saji & Imai 2013). We chose the 5 Korean verbs instead of the 7 Japanese or 13 Mandarin verbs described by Saji and Imai (2013) to limit the memory burden. The real Korean meanings varied in their breadth and overlap, but we created meanings of approximately the same scope per verb and without overlap. Each verb encoded a manner of contact or support (*antta*: support on shoulder or back; *ida*: support on one or two palms; *kidda*: hug or clutch against front or side; *meda*: dangle from arm, finger, or hand; *teulda*: grasp or grip with one or two hands; see Figure 1). Each of the five novel verbs was shown once as stationary and once with forward movement with each of four objects (yielding 8 video examples per verb and 40 videos total).

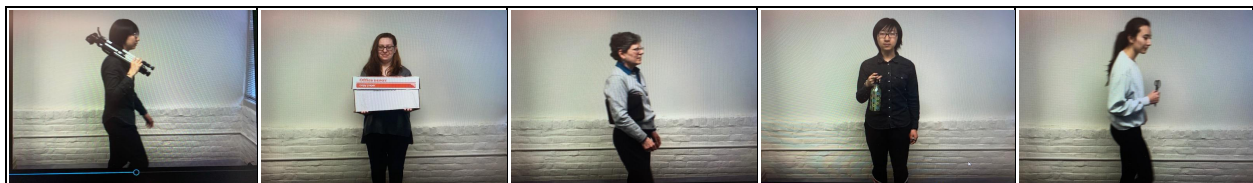


Figure 1: Examples of the five verbs to be learned. From left: Antta, Ida, Kidda, Meda, Teulda

Test stimuli were new instances of the same five verbs (with four new objects per verb and modest variations in their placement), shown once as stationary and once with movement for each object (yielding 40 test videos).

2.3. PROCEDURE. Participants were told that they would see brief video clips of a person interacting with an object and that each one would be labeled with a Korean verb. The Learn Labels group was told further only that “You’ll try to learn the correct label for each type of action.” The Figure Out Meanings group was told to try to learn the correct label “by figuring out what the verbs mean.” They were also told “Note that verbs in Korean may have different meanings from the ones you would use in English. Keep an open mind about what the meanings of the labels might be.”

There were two learning phases and a test phase. In Learning Phase 1, participants first viewed all the training stimuli with labels once, and then wrote what they thought each label meant. In Learning Phase 2, they chose a label for each stimulus and received feedback with the correct label. They then wrote again what they thought each label meant. In the Test Phase, they saw each novel stimulus and chose a label for it from a list of the five.

3. Results. We first discuss differences between the groups and then consider implications of the overall performance level of both groups.

3.1. GROUP DIFFERENCES.

3.1.1. LATENCIES. If the Learn Labels group has more difficulty shifting semantic dimensions and inducing the meanings of the verbs relative to the Figure Out Meaning group, they might take longer to complete the study. They might also take longer to make label choices in the test phase. (There were no choices in Training Phase 1, and choice times were not recorded in Training Phase 2). The Learn Labels group was slightly but not significantly slower than the Figure Out Meanings group to complete the study ($t(76) = -1.18, p > .10$). The non-significance may be due to the fact that fixed-time portions of the study such as video exposures swamped the impact of performance

differences in overall time. However, as expected, this group was significantly slower to complete the test trials ($t(76) = -1.84, p < .05$). (See Table 1.)

Group	Study Completion Time	Test Trial Choices
Learn Labels	24.84 mins	3.46 sec
Figure Out Meanings	21.43 mins	3.10 sec

Table 1: Latencies

3.1.2. DEFINITIONS. If the Learn Labels group has more difficulty shifting semantic dimensions and inducing the verb meanings, their definitions should less well capture the intended meaning of the verbs, and especially so in Learning Phase 1. We scored definitions in two ways: First, based on whether the definitions captured the general idea that meanings were based on the manner of contact or support, and second, whether they captured the specific intended meaning. Each definition (for each method) was assigned a 1 if closely capturing the target meaning, 0.5 if partially capturing it, and 0 if not similar to the target. Scoring was carried out by the three authors and a research assistant, with blocks of responses arranged in random order so that scorers were blind to condition. Each response was scored by all four scorers and its mean score was obtained.

In Learning Phase 1, the Learn Labels group was slightly but not significantly worse than the Figure Out Meanings group at both producing definitions mentioning manner of contact or support and at capturing the specific intended meaning ($t(76) < 1$ for both; $ps > .10$). In Learning Phase 2, the two groups were virtually identical for both. (See Table 2.)

Group	General Meaning	Specific Meaning
Learning Phase 1		
Learn Labels	.67	.47
Figure Out Meanings	.72	.51
Learning Phase 2		
Learn Labels	.79	.58
Figure Out Meanings	.77	.57

Table 2: Mean Score, Definitions

3.1.3. LABEL CHOICES. If the Learn Labels group has more difficulty shifting semantic dimensions and inducing the meanings of the verbs relative to the Figure Out Meaning group, their choices of labels for the actions should be less accurate in both Learning Phase 2 and in the Test Phase. (No choices were made in Learning Phase 1). The groups did not differ in Learning Phase 2, and, contrary to expectation, the Learn Labels group did somewhat better in the Test Phase ($t(76) = -1.6, p = .06$; see Table 3).

3.2. OVERALL PERFORMANCE. Overall, both groups performed at high and similar levels and showed improvement over the course of exposure.

3.2.1. DEFINITIONS. As seen in Table 2, each group increased the extent to which their definitions captured the general idea of body contact or support from the 1st to the second round of definitions, and each group also increased the extent to which their definitions captured the specific intended meaning from the first to the second round of definitions. The changes were significant for the Learn Labels group ($t(70) = 1.84$ for general idea and $t(70) = 1.73$ for specific intended meaning, $p < .05$ for both) but not for the Figure Out Meanings Group, whose performance was higher to begin with (for both, $t(82) < 1.0, ps > .10$).

Group	% Correct
Learning Phase 2	
Learn Labels	.74
Figure Out Meanings	.73
Test Trials	
Learn Labels	.83
Figure Out Meanings	.75

Table 3: Label Choices

3.2.2. LABEL CHOICES. Because there were five labels to choose from for each stimulus, chance performance would be 20% correct. As seen in Table 3, performances of both groups at Learning Phase 2 and at the Test Phase were well above this level (tested against .20, all $t_s > 12$, all $p_s < .0001$), indicating that substantial learning took place.

4. Discussion. L2 learners often show difficulty reshaping L1 word meanings to L2 meanings. Is this difficulty inherent in the nature of lexical networks (due to L1 entrenchment and competition between mappings), or is it due to the nature of typical L2 learning conditions? Our study was designed to help discriminate between these possibilities. It provided intensive exposure to novel word-referent pairings, fostering abstraction of commonalities and contrasts among the words and facilitating a focus on word meaning extraction rather than on understanding a larger discourse. Half the participants were also cued to expect that meanings might differ from L1.

4.1. GROUP COMPARISON. Half of our participants were asked simply to learn the novel labels for the actions they viewed, and half were cued to look for meanings different from those of their L1. The Figure Out Meanings group was significantly faster to make choices in the final test. They also had a non-significant edge in producing correct definitions of the novel words after the first round of exposure. However, they did not have more correct choices of labels in the second learning phase or the test phase. Because there was no hint of advantage for the Figure Out Meanings group and the sample sizes were not small, it is unlikely that stronger evidence of choice differences would emerge with additional participants. The similar choice performance between groups indicates that learners can do well even without prior knowledge that they might need to attend to different dimensions. The fact that their native language provides only two labels for the actions (*carry* and *hold*) whereas they were asked to learn to discriminate among use of five different Korean labels may have been sufficient to alert both groups that they needed to approach the task looking for meanings different from *carry* and *hold*. In real world L2 exposure, learners may not encounter the complete vocabulary for a domain within a single context, making it much less obvious that meaning differences must exist.

4.2. OVERALL PERFORMANCE. Both groups produced levels of choice performance well above chance in the test phase, despite the short study duration. This outcome indicates that it was not difficult to transition from their L1 system to a different system of labeling, again favoring the possibility that it is typical conditions of learning rather than inherent properties of lexical networks that create difficulty in real world L2 acquisition.

However, it should be noted that performance on the final test did not exceed 83% correct, and it will be important to examine the nature of the errors in future analyses, in conjunction with examining errors in definitions produced at the second learning phase. It was apparent in scoring the definitions that learners sometimes simply confused which label went with which meaning

(although they had induced relevant meanings). There was also some evidence in the definitions that a minority of participants did have difficulty making the meaning transition, persisting in producing definitions that referred to moving or standing still with the objects. It remains to be determined what the bulk of label choice errors in the test phase consist of, and whether those errors point to a particular source of difficulty in transitioning from English meanings to the novel verb meanings.

4.3. ADDITIONAL CONSIDERATIONS. Aside from input conditions, the task given to our participants was easier than the task that would be facing English learners of real Korean (or Mandarin or Japanese) in several respects. We designed our novel verb meanings to have approximately equal scope and no overlap among the meanings. In reality, these languages have terms with fuzzy boundaries and some that are broad in scope and others that are more specific. We also limited the number of verbs to five, whereas some languages have up to about 13 terms regularly used within the domain (Saji & Imai 2013). Such variations are likely to provide extra challenges to the learner in terms of figuring out the meanings associated with each term. However, although inducing the correct meanings may be more difficult under these circumstances, understanding that the nature of the meanings must fundamentally differ from that of the English system should, if anything, be even more clear.

Another consideration is that this work was prompted by observations of difficulty by L1 Mandarin speakers in grasping English *carry* and *hold* (Malt, 2020), but (to obtain larger sample sizes) the current study taught L1 English speakers meanings modeled on the Asian languages. Thus, the participants were moving from a simpler, two-word system to a more complex five-word system. Would learning be different if we had asked speakers of an Asian language to learn English meanings in the lab task, where they would need to collapse multiple meanings to fewer instead of vice versa, as well as make a dimensional shift? Previous work on learning in natural contexts suggests that going from more to fewer meanings is not necessarily easier than going from few to many (e.g., Jessen & Cadierno 2013, Gathercole & Moawad 2010). Ultimately, however, this is an empirical question, and we are currently preparing to run L1 Mandarin speakers in China in such a task.

Another difference between moving from L1 English to L2 Korean-like meanings compared to vice versa is that the Korean-like meanings were based on different manners of interacting with objects, whereas the English *carry-hold* distinction is based on presence or absence of movement. Even infants are sensitive to both the manner and path of events (Pulverman, Song, Hirsh-Pasek, Pruden & Golinkoff 2013). However, some evidence suggests that they may be more sensitive to path (Konishi, Pruden, Golinkoff & Hirsh-Pasek 2016). To the extent that the English *carry-hold* distinction is similar to a path distinction, its salience could provide a boost in identifying the relevant underlying dimension for those starting with Korean-like (or real Korean or Japanese or Mandarin) meanings. However, the current study also makes clear that noticing the different manners was not difficult under favorable conditions. Again, an empirical test is needed to establish if there is a difference in ease of moving from one system to the other.

Finally, although the current results suggest that learners can shift to new meanings with suitable input, an open question is whether such shifts will be accompanied by alterations to the L1 word representations. Some prior research points to the possibility that L2 attainment does exert an influence on L1 form-to-meaning mappings (e.g., Malt et al. 2015, Pavlenko & Jarvis 2002), perhaps by influencing connection strengths between forms and meanings (Malt et al. 2015). If so, it will be of interest to pinpoint how properties of lexical networks interact with language input in maintaining or altering representations in the two-language case.

4.4. CONCLUSIONS. High levels of choice performance in our task indicate that most learners can readily pick up on new dimensions of L2 word meaning that cross-cut their familiar L1 meanings. Success in the current task context is likely due to viewing many instances of word-referent pairings in succession (fostering abstraction of commonalities per word and identification of contrasts among the words) and attention dedicated to the learning task rather than to other tasks that may be more important in real-world contexts. As such, the results point to input conditions rather than inherent conditions of lexical networks as a critical limiting factor in L2 word learning.

References

- Barsalou, Lawrence. 1991. Deriving categories to achieve goals. *Psychology of Learning & Motivation* 27. 1–64.
- Bowerman, Melissa. 1996. Learning how to structure space for language: A cross-linguistic perspective. In Paul Bloom, Mary A. Peterson, Lynn Nadel, and Merrill F. Garrett (eds.), *Language and space*, 385-436. Cambridge, MA: MIT Press.
- Clark, Eve. V. 1987. The principle of contrast: A constraint on language acquisition. In Brian MacWhinney (ed.), *Mechanisms of language acquisition*, 1–33. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Groot, Annette. M. B. de. 2012. Vocabulary learning in bilingual first-language acquisition and late second-language learning. In Miriam Faust (ed.), *The Handbook of the Neuropsychology of Language*, vol 1, 472-493. Wiley-Blackwell, Oxford: UK. doi: 10.1002/9781118432501.ch23
- Gathercole, Virginia C. M. & Ruba A. Moawad. 2010. Semantic interaction in early and late bilinguals: All words are not created equally. *Bilingualism: Language and Cognition* 13. 385-408.
- Jessen, Moiken, & Teresa Cadierno. 2013. Variation in the categorization of motion in L2 Danish by German and Turkish native speakers. In Juliana Goschler & Anatol Stefanowitsch (eds.), *Variation and change in the encoding of motion events*, 133–162. Amsterdam: John Benjamins.
- Konishi, Haruka, Shannon M. Pruden, Roberta M. Golinkoff & Kathy Hirsh-Pasek. 2016. Categorization of dynamic realistic motion events: Infants form categories of path before manner. *Journal of Experimental Child Psychology* 152. 54-70.
- Kruschke, John K. 1996. Dimensional relevance shifts in category learning. *Connection Science* 8. 225-247.
- Malt, Barbara C. 2020. Understanding L2 word learning outcomes: The roles of semantic relations, input, and language dissimilarity. *International Journal of Bilingualism* 24. 478–491.
- Malt, Barbara C., Ping Li, Aneta Pavlenko, Huichun Zhu & Eef Ameel. 2015. Bidirectional lexical interaction in late immersed Mandarin-English bilinguals. *Journal of Memory and Language* 82. 86-104.
- Malt, Barbara C. and Asifa Majid. 2013. How thought is mapped onto words. *Wiley Interdisciplinary Reviews in Cognitive Science* 4. 583-597. doi: 10.1002/wcs.1251.
- Malt, Barbara. C. and Steven A. Sloman. 2003. Linguistic diversity and object naming by non-native speakers of English. *Bilingualism: Language and Cognition* 6. 47-67.
- Pavlenko, Aneta. 2009. Conceptual representation in the bilingual lexicon and second language vocabulary learning. In Aneta Pavlenko (ed.), *The bilingual mental lexicon: Interdisciplinary approaches*, 125 – 160. Bristol, U.K.: Multilingual Matters.

- Pavlenko, Aneta & Scott Jarvis. 2002. Bidirectional transfer. *Applied Linguistics* 23. 190-214.
- Pulverman, Rachel, Lulu Song, Kathy Hirsh-Pasek, Shannon M. Pruden & Roberta M. Golinkoff. 2013. *Child Development* 84. 241-252.
- Ross, Brian H. and Gregory L. Murphy. 1999. Food for thought: Cross-classification and category organization in a complex real-world domain. *Cognitive Psychology* 38. 495–553.
- Saji, Noburo, Mutsumi Imai, Henrik Saalbach, Yuping Zhang, Hua Shu & Hiroyuki Okada. 2011. Word learning does not end at fast-mapping: Evolution of verb meanings through reorganization of an entire semantic domain. *Cognition* 118. 45–61.
- Saji, Noburo & Mutsumi Imai. 2013. Evolution of verb meanings in children and L2 adult learners through reorganization of an entire semantic domain: The case of Chinese carry/hold verbs. *Scientific Studies of Reading* 17. 71-88.
- Snow, Catherine E. 1972. Mothers' speech to children learning language. *Child Development* 43. 549–565.