Research Article

SEMANTIC AND STRUCTURAL TASKS FOR THE MAPPING COMPONENT OF L2 VOCABULARY LEARNING

TESTING THE TOPRA MODEL FROM A NEW ANGLE

Shusaku Kida*

Hiroshima University

Joe Barcroft

Washington University in St. Louis

Abstract

The type of processing–resource allocation (TOPRA) model predicts that increasing one type of processing (semantic, structural, or mapping oriented) can decrease other types of processing and their learning counterparts. This study examined how semantic and structural tasks affect the mapping component of second language (L2) vocabulary learning. Japanese-speaking L2 English learners attempted to map secondary meanings of 24 English homographs. Each participant studied them (a) while making pleasantness ratings about word meaning (mapping plus semantic processing); (b) while counting letters in each word (mapping plus structural processing); and (c) without any additional task (mapping only). Results of L1 (first language) and L2 free recalls and L2-to-L1 and L1-to-L2 cued recalls indicated higher free recall in the semantic condition over the structural condition and higher cued recall in the mapping condition over the semantic and structural conditions, providing qualitatively new evidence for TOPRA model predictions.

One defining characteristic of SLA is that it is multicomponential process. SLA involves the acquisition of not only multiple linguistic subsystems—phonology, lexis, (morpho) syntax, pragmatics, and so forth—but also numerous dissociable (separable) components within these subsystems. In (morpho)syntax, one pair of dissociable components of general interest is that of (a) explicit learning *about* a target structure versus (b) successful implicit learning *of* the structure (*learning* vs. *acquisition* in Krashenian terms). As Krashen (1985) has elucidated, the former is not the same as the latter and should not be expected to produce the latter (see also Schwartz, 1993, on *learned linguistic*

*Correspondence concerning this article should be addressed to Shusaku Kida, Institute for Foreign Language Research and Education, 1-7-1, Kagamiyama, Higashi-Hiroshima, Japan, 739-8521. E-mail: skida@hiroshima-u. ac.jp

knowledge vs. *linguistic competence*). A second language (L2) learner may gain a large amount of explicit knowledge about how gender agreement works in a particular L2, for example, but that type of knowledge in no way ensures successful implicit learning (acquisition) of the grammatical gender system of the L2 in question.

In contrast to (morpho)syntax, L2 vocabulary (lexis) involves a type of form-meaning mapping that generally precludes explicit rule learning. If an L2 learner intentionally studies a set of target vocabulary items, there is no form-function-oriented rule to be learned in the sense that a rule can be learned *about* a morphosyntactic structure. Nevertheless, there are several pivotal questions that we can ask about the multi-componential nature of vocabulary learning: Which components of vocabulary knowledge are dissociable from one another? When a learner is exposed to target vocabulary in the input, how do increases in processing for one component affect processing for and learning of other components? To what extent are word form, word meaning, and form-meaning mapping dissociable from one another? How do increases in processing for one of these three affect processing and learning of the others?

Investigating the relative effects of different types of processing on different components of word knowledge is critical to improving our understanding of L2 vocabulary learning. Much of the existing research has focused on the effects of a various training and input-presentation methods without systematically isolating separate components of word knowledge. Research has assessed the keyword method (e.g., Ellis & Beaton, 1993; Wang, Thomas, & Quellette, 1992), sentential contexts (e.g., Hulstijn, 1992; Prince, 1996), visual aids such as pictures (e.g., Comesaña, Perea, Piñeiro, & Fraga, 2009; Lotto & de Groot, 1998), and semantic clustering (e.g., Tinkham, 1993; Waring, 1997). Studies in these areas have advanced our understanding of the relative efficacy of different types of training and presentation methods, but their focus has not been on distinguishing specific components such as word form, word meaning, and the mapping component of L2 vocabulary learning.

A different body of research (e.g., Barcroft, 2002, 2003, 2004; Wong & Pyun, 2012) has focused more on the dissociable subcomponents of L2 vocabulary learning, assessing how increases and decreases in one type of processing affects others. In doing so, they have assessed predictions the type of processing–resource allocation (TOPRA) model (Barcroft, 2002, 2003), which asserts that if overall processing demands are sufficiently high, increases in semantic processing, structural processing, or both should decrease available resources needed to process for the mapping component of L2 vocabulary learning. It has been demonstrated, for example, that increased semantic processing associated with a task such as sentence writing can decrease L2 word learning substantially (Barcroft, 2004; Wong & Pyun, 2012), however, virtually no research in this area has focused specifically on the effects of different types of processing and the *mapping* component of L2 vocabulary learning.

To test TOPRA predictions related the mapping component of L2 vocabulary learning, the present study employed a novel research methodology involving a learning task directed at isolating the mapping component of vocabulary learning. We refer to the task as *alternate meaning mapping*, which is when a learner maps alternate meanings of known L2 word forms—which may be homonyms, homographs, or both—to meanings already known in the L1 (and for which there are L1 translations). Successful alternate meaning mapping involves learning no new word form (L2 and L1 word forms are

already known) and no new word meaning (primary meanings in L1 and L2 and secondary meanings in L1 are already known). Instead, it depends on the mapping component of vocabulary learning: mapping a known L2 word form to a known meaning, a meaning that corresponds to an L1 translation of the secondary meaning of the word (e.g., homonym or homograph) in question.

Setting aside the potential value of alternate meaning mapping in vocabulary research in general (a topic discussed later), in the present study, we asked Japanese-speaking learners of L2 English to attempt to learn secondary meanings of English homographs that were also homonyms, such as when learning that the word *foot* means not only 足 (pronounced as *ashi*, as in *hand and foot*) but also 支払う (pronounced *shiharau*, as in *foot the bill*).[1] Vocabulary learning of this nature depended in particular on the mapping component of vocabulary knowledge because no new word forms nor word meanings were the target of the learning task. Learning an alternative meaning of an L2 (English) word such as *foot* does not involve relearning the word form in question nor does it involve relearning the L1 (Japanese) meaning 支払う (*foot* as in *foot the bill*).[2] What it does involve, critically, is mapping the known form *foot* on to the known meaning 支払う (*foot* as in *foot the bill*). As such, this learning task allowed us to test TOPRA predictions about the effects of increases in semantic or structural processing on the mapping component of L2 vocabulary learning.[3]

LEXICAL INPUT PROCESSING AND THE EFFECTS OF SEMANTIC AND STRUCTURAL TASKS

As with other areas of language acquisition, vocabulary learning depends on exposure to input (samples of the target language) and input processing (IP). Following Barcroft (2015), *lexical IP* (lex-IP), or "the manner in which learners process words and lexical phrases as input" (p. 14), is one of many different levels of IP engaged during language acquisition. Others include syntactic IP, pragmatic IP, and discourse IP within the larger overall process of *multilevel* IP. One way of improving our understanding of lex-IP is to conduct research on the effects of different tasks and ways of presenting target vocabulary in the input. Studies by Royer (1973), McNamara and Healy (1995), Barcroft (2007), and Harrington and Jiang (2013) have demonstrated, for example, that presentation of target words in a manner that provides learners with opportunities to retrieve them (or attempt to retrieve them) led to increased L2 vocabulary learning when compared to conditions that do not favor or involve retrieval. These findings, which are consistent with the findings of an array of research on the benefits of retrieval in human memory (see, e.g., Roediger & Guynn, 1996), suggest that the act of retrieving a word, either partially or fully, modifies developing lexical representations and promotes their development during lex-IP. Other studies have focused on the effects of different semantic and structural tasks, oftentimes in consideration of the levels-of-processing (LOP) framework for research on human memory (Craik & Lockhart, 1972) as it may relate to both first language (L1) and L2 vocabulary learning (e.g., for L1, Pressley, Levin, & Miller, 1982a; Pressley, Levin, Kuiper, Bryant, & Michener, 1982b; for L2, Coomber, Ramstad, and Sheets, 1986; San Mateo Valdehíta, 2013).

According to the LOP framework, human memory trace (memory strength) can be described as a function of the depth of information processing. Greater depth or deeper

processing, which "implies a greater degree of semantic or cognitive analysis" (Craik and Lockhart, 1972, p. 675), results in better memory performance than does less depth, or shallower processing. Consistent with LOP predictions, Hyde and Jenkins (1969) demonstrated that performing a semantically oriented task, such as making pleasantness rating about words, can lead to better recall performance of known L1 words when compared to structurally oriented tasks, such as estimating the number of letters in words. As another example, Craik and Tulving (1975) reported that when participants were asked to perform semantic processing tasks such as semantic categorization (e.g., participants judged whether the word in question was a member of a category by answering a question such as "Is the word a type of fish?") or sentence completion (e.g., participants judged whether the word in question fit a sentence context), their subsequent memory performance was better as compared to when they were asked to perform structural processing tasks, such as a typescript task (i.e., participants judged if the word presented was upper-/lowercase) or a rhyming task (i.e., participants judged if the word rhymed with another presented word). Findings of this nature speak to the potential superiority of semantic elaboration over structural elaboration when it comes to memory for known forms and, in particular, known (previously acquired) L1 words.

Morris, Bransford, and Franks (1977) proposed transfer-appropriate processing (TAP) as an alternative framework for interpreting the benefits of semantic processing over structural in addition to learning contexts in which structural processing yields benefits over semantic specific processing. According to TAP, semantic processing does not always lead to better memory performance when compared to structural processing. Instead, the relative effectiveness of a task should be determined by the degree of similarity in the type of processing engaged by a task at study and the type of processing required for successful performance of a task at test. In their seminal study, Morris et al. demonstrated that when a standard recognition test (i.e., target items were original to-be-remembered words) was used, a semantic task (e.g., sentence-completion) led to better performance when compared to a structural processing task (e.g., a rhyming), a typical LOP effect. When a structurally oriented test (i.e., target items were rhymes of the original to-be-remembered words) was used, however, the structural task led to better performance as compared to the semantic task.

Similarly, Stein (1978) compared the effects of a semantic processing task (participants judged whether a preannounced meaning was expressed by each presented word) with a structural processing task (participants judged whether a preannounced letter was included in each presented word). Both semantic and structurally oriented recognition tests were administered. In the former test, participants were exposed to the target items and distractor items that were similar in meaning but different in form. In the latter test, they were exposed to the target items and distractor items that were different in meaning but similar in form. Results indicated that memory performance in the semantic processing condition was better than in the structural processing condition when the participants took the semantically oriented recognition test; however, the opposite was the case for the structurally oriented test.

Considering various demonstrations of the positive effects of semantic elaboration and increased semantic processing ("deeper processing" from the LOP perspective) on memory for known L1 words, many researchers became interested in whether semantic tasks might be used to facilitate L2 vocabulary learning. The evidence favoring the use of

semantic tasks and "deeper" semantically oriented processing to facilitate L2 vocabulary learning is extremely limited (if not nonexistent), however. The bulk of research in this area suggests that semantic tasks can impede L2 vocabulary learning, particularly the formal component, when overall processing demands are sufficiently high (as discussed later). Nevertheless, let us consider two cases in which claims have been made about purported benefits of semantically oriented processing on pseudoword learning (as a proxy for new word learning) and L2 word learning.

As a first case, Coomber et al. (1986) asked participants to learn 10 pseudowords while comparing the effects of three types of rehearsal: (a) definitions, (b) usage examples, and (c) writing novel sentences. Three types of posttests that corresponded to each type of rehearsal administered (match pseudowords to a definition, match pseudowords to a usage example, and write pseudowords in sentences). Their results demonstrated positive effects of sentence writing over the definition method, which led Coomber et al. to conclude empirical support for the LOP framework in the realm of L2 vocabulary learning. As a second case, San Mateo Valdehíta (2013) compared three learning methods somewhat similar to those used by Coomber et al: (a) choose the right definition for each target word, (b) choose an example that includes a semantically equivalent word for each target word, and (c) write a sentence that includes each target word. Consistent with the results of Coomber et al., San Mateo Valdehíta demonstrated positive effects of the sentence writing methods over the other two methods.

These two studies exemplify research in which specific tasks deemed to involve more semantic processing led to better novel word (or pseudoword) learning when compared to other tasks deemed to involve less semantic processing. As such, one might argue that they support the applicability of LOP to L2 new word learning. However, the assumption that the sentence writing was the most semantically oriented (or "deeper") of the tasks in these studies is questionable given that tasks such as choosing and analyzing one appropriate definition (out of several) ultimately may involve more semantically oriented processing than sentence writing (see Barcroft, 2015, for more on this possible interpretation of the results of comparisons of the effects of multiple types of semantic tasks of this nature).

In sum, the extension of the LOP framework into word learning has not been sufficiently qualified. Much of previous research assessed the relative veracity of the idea that increases in semantically oriented processing should lead to improved L2 vocabulary learning, however, the combined results of these studies did not provide any convincing support for extending an unqualified extension of LOP to L2 vocabulary learning (see Barcroft, 2002, for a review). Therefore, alternative approaches should, at minimum, be entertained to capture the effect of different types of processing on the acquisition of L2 vocabulary.

THE TOPRA MODEL: AN ALTERNATIVE TO UNQUALIFIED EXTENSIONS OF LOP

Considering lack of evidence regarding any unqualified application of LOP to L2 vocabulary learning, Barcroft (2002, 2003) proposed the TOPRA model (Figure 1) as a means of assessing the impact of semantic, structural, and mapping-oriented processing on lex-IP and vocabulary learning. The two thick outer bars in all versions of the TOPRA model do not move because they reflect overall limits on a learner's processing capacity at any given point in time. As depicted in Figure 1A, according to the TOPRA model, increases in one type of processing, such as Processing Type A, can decrease processing
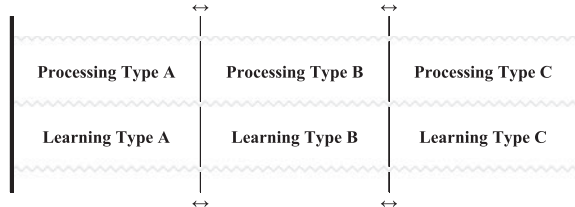
resources available for other types of processing, such as Processing Types B and C. The result of such interaction is a trade-off, reflected in the corresponding learning counterparts for A, B, and C. This trade-off concerns any set of dissociable types of processing in the general version of the model (Figure 1A) and in specific terms regarding different components of word learning in the second version of the model (Figure 1B). The third version of the model (Figure 1C) is designed to focus specifically on two types of processing related to two specific components of vocabulary learning, in this case the semantic and formal components.

TOWARD MORE FINE-GRAINED ASSESSMENTS OF L2 VOCABULARY LEARNING
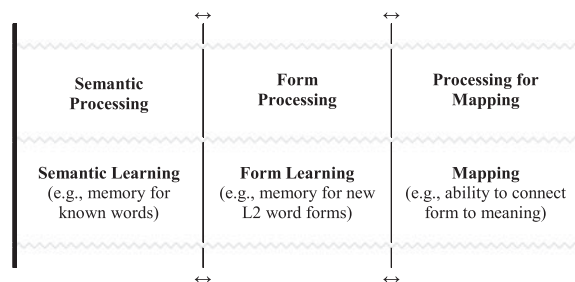
Even though L2 vocabulary acquisition involves multiple components of knowledge (see, e.g., Nation, 2001, 2007), studies oftentimes have assessed only one component or
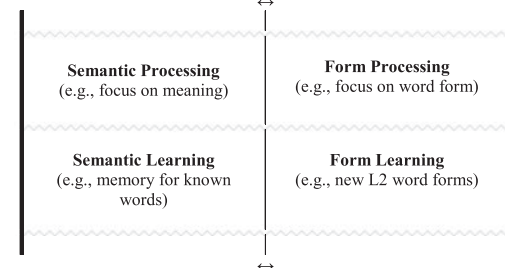
FIGURE 1.   The TOPRA model (Barcroft, 2002, 2003).

aspect of this knowledge (Waring & Takaki, 2003). Relatedly, because TOPRA makes predictions about how different types of processing affect different components of vocabulary knowledge, studies designed to test TOPRA predictions need to consider the extent to which the assessment measures they include are dependent upon one or more different components of vocabulary knowledge. Therefore, before reviewing previous studies on the TOPRA model, let us consider Table 1, which summarizes how different components of vocabulary knowledge are assessed by different tasks that might be used as posttest measures.

For L1 free recall, performance depends largely on semantic activation within L1 and not on L2 word form retrieval. For L2 free recall, however, performance depends largely on one's degree of learning and ability to retrieve L2 word forms only (graphemes and their sequences in the writing mode and phonemes and their sequences in the spoken mode). Mapping is not critical to L2 free recall because one need not remember what the L2 word forms in question mean to retrieve and produce them. This is not the case with L1-to-L2 and L2-to-L1 cued recall because they depend on the degree to which one has learned L2 word forms as well as mappings between the L2 word forms and their meanings.

For L1-to-L2 cued recall, performance depends particularly heavily on the degree to which one has learned L2 word forms because it requires learners to produce the L2 word forms on their own (the cue being a picture or an L1 translation) whereas L2-to-L1 does not (the cue being the L2 word form). In other words, the direction of cued recall makes a difference in terms of the sensitivity of learners' knowledge about the L2 word form in question. Learners need to have more knowledge of the form to produce correct responses for picture-to-L2 or L1-to-L2 cued recall while they do not have to have complete knowledge for L2-to-L1 cued recall (see, e.g., Barcroft, 2009). Furthermore, it is difficult to determine the extent which performance on picture-to-L2 cued recall, L1-to-L2 cued recall, and L2-to-L1 cued recall depend on having learned the mapping component of vocabulary learning. Performance certainly depends on knowing the

TABLE 1.   Recall types and the type of memory/learning on which each depends

Recall Type	It depends largely on …
For a known L1 word …	
(1) L1 free recall	semantic activation within L1 lexicosemantic space.
For a novel L2 word form with a known L1 meaning …	
(2) L2 free recall	learning L2 word form.
(3) L1 (or picture)-to-L2 cued recall	learning L2 word form and form-meaning mapping.
(4) L2-to-L1 cued recall	learning L2 word form (but to a lesser degree than 3) and form-meaning mapping.
For a novel L2 homograph with a known L2 word form and a known L1 meaning* …	
(5) L1-to-L2 cued recall	mapping known L2 words to known meaning.
(6) L2-to-L1 cued recall	mapping within L2 words to known meaning.

*Alternate meaning mapping
Note: This scheme does not negate the need to learn L2-specific meaning and usage (for 2–4) nor the possibility of L1 and L2 mutual activation/shared lexicosemantic space.

mapping component, but performance tends to depend heavily on having learned the formal component, especially in the case of picture-to-L2 or L1-to-L2 cued recall for which production of L2 word forms is required.

Finally, alternate meaning mapping—mapping a known L2 word onto a known meaning (conveyed by an L1 translation)—depends on the extent to which a learner has learned the mapping component of vocabulary knowledge. For alternate meaning mapping, performance does not reflect how well a word form or a word meaning has been learned because the word form and word meanings in question were already known prior to the time at which a learner attempts to do the alternate meaning mapping.

Given that the TOPRA model makes predictions about how different types of processing affect different components of vocabulary learning, it is useful for studies focused on TOPRA to consider how different assessment measures reflect learning of different components of vocabulary learning in varying degrees. Table 1 provides what we hope to be a useful summary for six distinct types of assessment measures. With the summary provided in Table 1 as a general background, in the next section we review several previous studies that tested predictions of the TOPRA model from different angles.

PREVIOUS RESEARCH ON THE TOPRA MODEL

Most research to date has focused on this prediction given what L2 learners typically need to learn (encode and retain) at the initial stages of vocabulary learning: a new word form and a mapping between the word form and a known L1-based meaning considering that adult L2 learners transfer semantic representations from their L1 (see, e.g., Jiang, 2000; Laufer, 1997; Ryan, 1997). Adult English-speaking learners of L2 Spanish do not relearn primary meanings of words such as *manzana* "apple," *puerta* "door," and *libro* "book," for example, when they encounter these words in the input. Even though they do need to refine L2-specific semantic space for all words (learning L2-specific connotative and denotative meanings, appropriate usage space, collocational properties, etc.), meanings with viable L1 counterparts already known are transferred when the L2 versions of these words are first learned.

These points being clarified, various studies (e.g., Barcroft, 2002, 2003, 2004; Wong & Pyun, 2012) have provided support for the semantic/structural bipartite version of the TOPRA model (Figure 1C) by demonstrating how semantically elaborative tasks can negatively affect L2 word-form learning when compared to no-task mapping conditions. Barcroft (2003) asked English-speaking university students enrolled in second-semester L2 Spanish to attempt to learn 24 novel words in Spanish (e.g., *borla* "tassel," *rastrillo* "rake," *embudo* "funnel") in two conditions. For 12 of the words, participants were instructed to address questions about word meaning (increased semantic processing condition) as follows: "When was the last time (if ever) that you used this object?" "In what ways can this object be used?" "Where can you buy this object?" "What other objects do you associate with this object?" For the other 12 words, they were instructed only to do their best to learn the new words and were not given any additional task (control condition). The posttest assessment measures were L2 and L1 free recall and picture-to-L2 and L2-to-L1 cued recall. Results indicated more free recall in L1 than in L2 but no effect of condition on either type of free recall. For both picture-to-L2 and

L2-to-L1 cued recall, by contrast, the questions condition resulted in lower scores. Picture-to-L2 cued recall in the question-answering condition was 5.03 as compared to 6.44 in the no-question-answering condition (maximum = 12), for example. These findings are consistent with the TOPRA model's prediction that increased semantic processing can decrease word form learning and, to the extent that it was a factor, the mapping component of L2 vocabulary learning as well.[4]

In two other studies, Barcroft (2004) and Wong and Pyun (2012) found that requiring learners to write target L2 words in sentences strongly diminished their intentional word learning when compared to no-task mapping conditions. Barcroft (2004) found that the sentence-writing condition decreased picture-to-L2 cued recall scores from 9.64 to 4.40 (Experiment 1; maximum = 12) and from 10.30 to 4.95 (Experiment 2; maximum = 12) as compared to no-sentence-writing conditions when English speakers attempted to learn L2 Spanish words. Wong and Pyun (2012) also demonstrated very strong negative effects for sentence writing on both picture-to-L2 and L2-to-L1 cued recall among English speakers learning L2 French and L2 Korean. Interestingly, the negative effects were more pronounced for L2 Korean, the language with forms that were less similar to the L1 (English) of the participants in the study, which is also consistent with TOPRA predictions.

In another study more closely related to the present study in terms of methodology, Barcroft (2002) assessed the effects of both semantic and structural tasks on intentional L2 vocabulary learning among English-speaking learners of L2 Spanish.[5] The semantic task was pleasantness ratings, asking participants to rate on a scale the extent to which the meaning of each word was pleasant or unpleasant to them. The structural task was letter counting, asking participants to count the number of letters in each word. These two tasks, also used in the present study, were selected considering their history in LOP research (e.g., Hyde & Jenkins, 1969). The participants attempted to learn eight words in each of three conditions: pleasantness ratings (semantic), letter counting (structural), and a no-task ("do your best" only) condition. After the study phase, all participants completed two free recall tasks—in L2 and in L1—and two cued recall tasks—picture to L2 and L2 to L1. The results indicated higher L1 free recall scores for the semantic task over the structural task, a typical LOP effect, but also higher L2 free recall scores for the structural task over the semantic task, an inverse LOP effect which, from the TOPRA perspective, is due to L2 free recall being dependent upon word *form* learning whereas L1 free recall was not.

THE CURRENT STUDY

One set of predictions of TOPRA that remains to be tested concerns the relationship between semantic processing, structural processing, and processing needed for the mapping component of L2 vocabulary learning. As is indicated in Figure 1B, TOPRA predicts that increases in semantic processing, structural processing, or both can reduce available processing resources for other types of processing and their learning counterparts, including the mapping component of vocabulary learning. From our perspective, the mapping-oriented learning involved in learning homographs (that are also homonyms) offers a unique opportunity in this regard. If L2 learners are asked to learn secondary meanings of homographs, provided that the secondary meanings already

exist in their L1, then the learning task in which they are engaging concerns primarily the mapping component of L2 vocabulary learning. If a Japanese-speaking learner of L2 English who knows the word form *dish* and its meaning in the sense of container but not its meaning in the sense of "to gossip" and knows what "to gossip" means in Japanese (having a translation for it and so forth), then the process of learning that the word *dish* can also mean "to gossip" depends largely (if not solely) on the mapping component of vocabulary learning for the word in question. Therefore, learning to map meanings of known L2 word forms that are homographs provides for a unique context in which to test the TOPRA model regarding processing and learning outcomes related to the mapping component of vocabulary learning.

Note also that homograph learning is distinct from many other types of L2 vocabulary acquisition. A more common type of L2 vocabulary learning is when learners must encode and retain new L2 word *form* and make new form-meaning mappings between the new form and its meaning. In such situations, it is also common for no new meaning to be learned during the initial stages because the meaning from L1 for the word is simply transferred (Jiang, 2000; Laufer, 1997; Ryan, 1997). Another type of vocabulary learning is learning cognates. Studies have found that cognate status has considerable effects on L2 word learning (De Groot & Keijzer, 2000; Ellis & Beaton, 1993; Lotto & de Groot, 1998), and as is probably intuitive, more familiar word forms that are easier to learn than less familiar ones. Homograph learning differs from cognate learning, however, in that for homograph learning, the word form in question is always identical. Additionally, homograph learning is also a unique type of vocabulary learning in L1 (Storkel & Maekawa, 2005) because whereas L1 word learning typically involves learning a new word form and meaning, homograph learning in L1 does not involve learning a new word form.

Also related to the task in the present study is learning polysemy in L2, which involves mapping known word forms to variant related word meanings. One example would be learning the word the *hand* as a noun and then learning that the same word form can mean *hand* as a verb. Most words are *polysemous*, having multiple meanings (Finkbeiner, Forster, Nicol, & Nakamura, 2004). Previous studies have tracked L2 learners' acquisition of peripheral meanings of known words longitudinally in naturalistic settings (Crossley, Salsbury, & McNamara, 2010; Schmitt, 1998) while others have examined effects of different types of polysemous word processing in more experimental settings (Bogaards, 2001; Verspoor & Lowie, 2003). Verspoor and Lowie (2003), for example, demonstrated that Dutch-speaking L2 English learners guess and retain target figurative meanings of polysemous words better when provided with target words in sentential contexts that contain *core* meaning as compared to *figurative* meaning (see also Morimoto & Loewen, 2007, on effects of image-schema-based vs. translation-based instruction of L2 polysemous words). Such research confirms that L2 learners can develop increasingly in-depth knowledge of the polysemous words by establishing new form-meaning mappings over time, but it has not been focused on the dissociability among different components of L2 vocabulary learning nor TOPRA-based predictions about relationships between these dissociable components.

In this light, the present study was designed to test predictions of the TOPRA model regarding effects of increases in semantic processing and structural processing (invoked by semantic and structurally oriented tasks) on the mapping component of L2 vocabulary

learning by asking Japanese-speaking L2 learners of English to attempt to map a series of alternate meanings of English words (homographs that were also homonyms) for which they already knew primary meanings. The three learning conditions were as follows:

- Semantic: Participants attempted to memorize L2 and L1 word pairs while performing a pleasantness-rating task, rating the extent to which the meaning of each word was pleasant or unpleasant to them by circling a number on a 5-point Likert scale.
- Structural: Participants attempted to memorize the L2 and L1 word pairs while performing a letter-counting task, counting the number of letters in each word and indicating responses by circling a number from 1 to 5.
- Mapping (Control): Participants attempted to memorize the L2 and L1 word pairs only.

The primary posttest measure of interest in the study was cued recall, that is, when participants were provided with each target L2 word and asked to map it onto a new L1 translation. The TOPRA model predicts that each of the two tasks in question (semantic and structural) can decrease processing for and learning of the mapping component of vocabulary learning by exhausting processing resources that otherwise could have allocated toward the mapping component. Free recall in L1 and L2 was also included to assess presence or absence of typical LOP effects in memory for known (L1) words and to explore other potential effects in this particular learning context.

Because learning to map secondary meanings of L2 homographs onto known L1 meanings does not involve learning novel word form or learning novel meanings, following the predictions of the TOPRA model, our hypotheses for cued recall were as follows.

(1) Increased semantic processing (pleasantness ratings) will not facilitate and may decrease performance.
(2) Increased structural processing (letter counting) will not facilitate and may decrease performance.
(3) The condition in which learners are not required to perform a semantic or structural task (mapping or no task) will result in the best performance.

Each of these three hypotheses would be supported in the present study if both the pleasantness-ratings and letter-counting tasks led to lower cued-recall performance.

METHOD

PARTICIPANTS

The participants in the present experiment were 112 Japanese-speaking L2 English learners from five intact classes.[6] All participants were first-year university students in Japan and were native speakers of Japanese. The mean English proficiency level of these participants was intermediate based on their scores of the Test of English for International Communication (TOEIC), which is a standardized English proficiency test developed by Educational Testing Service. The test measures English receptive competence in the multiple choice format (three or four options) consisting of a listening section (100 questions that include, e.g., picture descriptions and comprehension questions about dialogues or an individual's speech, lasting approximately 45 minutes) and a reading section (100 questions that include, e.g., fill-in-the-cloze type grammar or

vocabulary questions and reading comprehension questions about English passages, such as e-mails and ads, lasting about 75 minutes). The possible score of this test ranges from 10 to 990. The participants' TOEIC score ranged from 225 to 690. The mean was 468.93; the *SD* was 97.42.

EXPERIMENTAL WORDS

The target words consisted of 24 English words that were selected by referring to the General Service List (West, 1953) and an English polysemous word list (Durkin & Manning, 1989). The General Service List was used because a central goal of the present study was to find words whose form and primary meaning would already be known by the L2 learners, and many of the high-frequency words that appear in the General Service List satisfied this criterion. Of course, the same high-frequency words could have been obtained from another English frequency list, but the General Service List served our goals in this regard without complication. All the words were one-syllable whose word class for their primary meaning was determined to be noun, such as *foot* whose primary meaning translated to Japanese as 足 as in *hand and foot* and whose secondary meaning, a verb, translated as 支払う as in *to foot the bill*. All experimental words and their translations appear in the appendix.

DESIGN

Participants learned target words in each of three conditions: (a) make pleasantness ratings about each word's meaning (semantic); (b) count letters in each English word (structural); and (c) "do your best" only (mapping). The study included three learning orders, or *lists* (Figure 2). The participants were randomly assigned to one of the three list conditions (38 participants had list 1, 35 had list 2, and 39 had list 3). In list 1, words 1 to 8 (*head*, *bowl*, *shell*, *cow*, *trip*, *coast*, *floor*, *dog*) were learned in the semantic condition, words 9 to 16 (*bolt*, *plug*, *dish*, *man*, *bust*, *sock*, *stump*, *stem*) in the structural condition, and words 17 to 24 (*dress*, *foot*, *court*, *snow*, *line*, *ice*, *rock*, *case*) in the mapping condition. In list 2, words 1 to 8 were learned in the mapping condition, words 9 to 16 in the semantic condition, and words 17 to 24 in the structural condition. Finally, in list 3, words 1 to 8 were learned in the structural condition, words 9 to 16 in the mapping condition, and words 17 to 24 in the semantic condition. Therefore, each target word was shown with an equal amount of exposure in each list. Lexical characteristics were controlled among these three sets (words 1 to 8, words 9 to 16, and words 17 to 24).[7]

PROCEDURES

Data were collected in the participants' regular classrooms during regular class hours according to the following procedures.

(1) The participants, who were instructed that their participation in the experiment was voluntary, completed informed consent forms.
(2) The researcher administered a pretest, which required the participants to choose the primary meaning of each target word from four options (one correct and three distractors), as in the following case for the English word *foot*: 過労 (overwork), 魅力 (fascination), 足 (foot), and 圧力 (pressure). Additionally, they were asked to write all possible translations for each target

	Word 1 to 8	Word 9 to 16	Word 17 to 24
List 1	Semantic	Structural	Mapping
List 2	Mapping	Semantic	Structural
List 3	Structural	Mapping	Semantic

FIGURE 2.   Presentation orders and conditions.

word. The order of presentation was random so that it was completely different from the order of word presentation during the learning phase. We used a productively oriented cued recall (see English word, produce secondary meaning in Japanese) for the pretest rather than a recognition test involving the secondary meaning to avoid exposing the participants to the target secondary meanings of interest before the learning phase. Exposing the participants to the secondary meanings of interest on the pretest might have resulted in undesired testing effects.

(3) After the pretest, the participants were provided with the general instructions for the experiment. They were instructed to attempt to learn new meanings for English words they already knew with six seconds for each target word in each of three different learning conditions. The target items were presented on a screen at the front of the classroom using a presentation program on a laptop computer. The participants were also informed that they would be tested afterward. The specific instructions given for the semantic and structural conditions were almost the same as those used by Barcroft (2002). The specific instructions for the semantic condition were as follows: "Make pleasantness ratings about each item that you see. One is the lowest ranking, and 5 is the highest ranking. A ranking of 1 means that you find the item to be extremely unpleasant for some reason or another. A ranking of 5 means that you find the item to be extremely pleasant for some reason or another. Carefully make a pleasantness ranking for each word based on your experiences, general feelings, or both. Please do your best to learn these eight words." The specific instructions for the structural condition were as follows: "Count the number of letters in each word. Circle the number on the scale that corresponds to the correct number of letters for each word. For example, by circling the number of 4 for a word, you are indicating that you counted four letters in this word. Please do your best to learn these eight words." The specific instructions for the mapping condition were as follows: "Your only task is to do your best to learn the new words. Please do your best to learn these eight words" (Barcroft, 2002, p. 340).

(4) After receiving these instructions, each participant moved on to the practice session in which they were exposed to the three learning conditions with six words that were not used in the main experiment. Specific instructions for each learning condition were shown again. L2 words and their translations were shown on a screen in the classroom. The participants had 26 seconds to read the instructions before beginning the new learning phase.

(5) After the practice session, the main learning phase began. The procedure for the main learning phase was the same as in the practice session except that the participants had three trials in the same order.

(6) Immediately after the study phase, four posttests were administered in the following order: L1 free recall, L2 free recall, L2-to-L1 cued recall, and L1-to-L2 cued recall. We used this order in consideration of certain types of possible testing effects, namely, that the free recalls needed to be administered before the cued recalls to avoid providing the participants with additional exposures to the target words immediately before their attempt to do the free recalls. In the L1 free recall test, the participants were asked to recall and write as many Japanese translations as possible in any order within two minutes. In the L2 free recall test, they were

asked to recall and write as many English target words as possible in any order within two minutes. In the two subsequent cued recall tests, each cue (the L2 target word in the L2-to-L1 cued recall and the L1 Japanese translation in the L1-to-L2 cued recall) was shown on the screen for 12 seconds. The order of presentation of the cues on the screen in the two cued recall tests was completely random and did not have any relation to the order of word presentation in the study phase.

Note that all the experimental procedures described in the preceding text (i.e., pretest, alternate meaning mapping, and four types of recall) took place on the same day, making it possible that the act of taking one test affected performance on the others. Critically, however, the order of the test battery, along with ordering of word groups and conditions (following a Latin square design) did not favor any one of the three conditions over any other.

SCORING

For each of the four posttests, binary scoring was adopted, that is, 1 point was assigned for each correct answer, and 0 was assigned for each incorrect answer. No partial points were given for other responses because the participants already knew the target word forms and their primary meanings based on their experience in L1 (Japanese). The element that was new for them was the need to map the L1-based meanings onto already known L2 forms. For example, participants already knew the English word *foot* and its primary meaning translated as 足, as in *hand and foot*. They also knew the Japanese word 支払う, the translation for the secondary meaning of the word *foot*, as in *foot the bill*. However, they did not know that the word *foot* has the secondary meaning of 支払う. Therefore, successfully making the connection between an English word and its secondary meaning on the cued recall posttests was an all-or-nothing proposition, making any type of partial scores unnecessary. Further, preknowledge of primary and secondary meaning was also scored in a binary manner (i.e., 1 was assigned if the participant knew the primary and secondary meaning of the target words *before* the experiment, and 0 was assigned if they did not).

STATISTICAL ANALYSES

In the present study, we conducted logit mixed-effects analyses on the four sets of recall data separately using the *lme4* package (Bates, Maechler, Bolker, & Walker, 2015; R Development Core Team, 2013) with participants and items as cross-random factors. Mixed-effects models have several advantages over conventional analyses of variance (ANOVAs) (Linck & Cunnings, 2015). They are more flexible when it comes to data distribution and allow researchers to analyze data scored in a binary manner (e.g., correct or incorrect, as in the present study) that are not able to be analyzed using traditional parametric statistics that require the data to be normally distributed. Mixed-effects models are also flexible when it comes to handling data with missing values. Therefore, they were more desirable for present purposes when compared to an ANOVA.

In the analyses for the logit mixed-effects models, we adopted both forward and backward model selection procedures. We used the *bobyqa* optimizer in lme4 to avoid convergence failure as much as possible. In the backward model selection, we first constructed a maximal model for each test result separately (L1 free recall, L2 free recall,

L2-to-L1 cued recall, L1-to-L2 cued recall). For each recall result, we included condition (semantic, structural, mapping), L2 proficiency based on TOEIC score (transformed as z score and then centered), and preknowledge of primary and secondary meanings of target words as fixed effects. Interactions among all variables were included to detect any possible interaction. We also included random intercepts for participants and items, by-participants' random slopes for condition; preknowledge of primary and secondary meanings; and by-item random slopes for condition, L2 proficiency, and preknowledge of primary and secondary meanings. After observing that the maximal model did not converge, we removed random slopes of variables one by one until the model converged.

In the forward model selection, we first constructed the unconditional model that included only by-participant and by-item random intercepts. We then included each variable in a step-by-step manner. The final model was determined based on the results of the likelihood ratio test and information criteria (Akaike Information Criterion and Bayesian Information Criterion). The two procedures (forward and backward) identified the same model for all four sets of recall results. The final four models included only condition as a fixed effect variable and random intercepts for both participants and items.

We report effect sizes for the four final models. One recently proposed way to report effect sizes for generalized linear mixed-effects models is to report both marginal R^2 and conditional R^2 values (Nakagawa & Schielzeth, 2013). The former concerns variance explained by fixed factors; the latter variance is explained by both fixed and random factors (Nakagawa & Schielzeth, 2013). We report the values by using the *MuMIn* package (Bartoñ, 2015). Finally, multiple comparisons were conducted with Tukey's honestly significant difference method using the *multcomp* package (Hothorn, Bretz, & Westfall, 2008).

RESULTS

Percentages for correct and incorrect responses for cued recall and free recall based on condition (semantic, structural, mapping) appear in Table 2. The results indicated the following: (a) for L1 free recall, the recall rate in semantic condition was higher than the recall rates in the other two conditions; (b) for L2 free recall, recall rates in the semantic and structural conditions were lower than the recall rate in the mapping condition; (c) for L2-to-L1 cued recall, recall rates in the semantic and structural conditions were lower than the recall rate in the mapping condition; (d) and finally, for L1-to-L2 recall, recall rates in the semantic and structural conditions were lower than the recall rate in the mapping condition.

Free Recall

Results of the mixed-effects models analyses appear in Table 3. Results for both L1 and L2 free recall indicated that condition was a significant factor (marginal $R^2 = 0.02$ and conditional $R^2 = 0.15$ for L1 free recall while marginal $R^2 < 0.01$ and conditional $R^2 = 0.13$ for L2 free recall). Results of multiple comparisons appear in Table 4.

Cued Recall

Results of the mixed-effects models analyses appear in Table 5. Again, results for both L2-to-L1 and L1-to-L2 cued recall indicated that condition was a significant factor

TABLE 2.   Responses, percentile scores, and 95% confidence intervals (CIs) in all test types under all conditions

	Correct	%	95% CI	Incorrect	%	95% CI
L1 free recall						
Semantic processing	453	50.56	[47.23, 53.88]	443	49.44	[46.12, 52.77]
Structural processing	332	37.05	[33.88, 40.31]	564	62.95	[59.69, 66.12]
Mapping	425	47.43	[44.12, 50.76]	471	52.57	[49.24, 55.88]
L2 free recall						
Semantic processing	425	47.43	[44.12, 50.76]	471	52.57	[49.24, 55.88]
Structural processing	423	47.21	[43.90, 50.54]	473	52.79	[49.46, 56.10]
Mapping	472	52.68	[49.34, 55.99]	424	47.32	[44.01, 50.65]
L2-to-L1 cued recall						
Semantic processing	556	62.05	[58.78, 65.24]	340	37.95	[34.76, 41.22]
Structural processing	521	58.15	[54.84, 61.40]	375	41.85	[38.60, 45.16]
Mapping	683	76.23	[73.30, 78.98]	213	23.77	[21.02, 26.70]
L1-to-L2 cued recall						
Semantic processing	676	75.45	[72.49, 78.23]	220	24.55	[21.77, 27.51]
Structural processing	690	77.01	[74.11, 79.73]	206	22.99	[20.27, 25.89]
Mapping	773	86.27	[83.84, 88.46]	123	13.73	[11.54, 16.16]

Note: Maximum score was 896 (number of items under each condition [8] × number of participants in each condition [112]).

(marginal $R^2 = 0.04$ and conditional $R^2 = 0.39$ for L2-to-L1 cued recall while marginal $R^2 = 0.03$ and conditional $R^2 = 0.47$ for L1-to-L2 cued recall). The results of multiple comparisons appear in Table 6. Given that the primary focus of this study was the mapping component of vocabulary learning reflected in these two recall tests, note that they show a clear advantage for the mapping condition over the other two conditions.

SUMMARY OF RESULTS

Overall, the free recall results indicated better memory performance in the semantic condition as compared to the structural condition, reflecting a typical LOP effect (semantic leading to better memory than structural processing). Of greater interest to the goals of the

TABLE 3.   Results of logit mixed-effects model for L1 and L2 free recall

Fixed Predictor	Estimate	SE	z
L1 Free Recall			
Intercept	−0.11	0.15	−0.73
Condition (Semantic vs. Mapping)	0.16	0.10	0.10
Condition (Structural vs. Mapping)	−0.46	0.11	<.01**
L2 Free Recall			
Intercept	0.13	0.15	0.87
Condition (Semantic vs. Mapping)	−0.22	0.10	−2.22*
Condition (Structural vs. Mapping)	−0.23	0.10	−2.32*

Note: SE refers to standard error; * $p < .05$, ** $p < .001$.

TABLE 4.   Multiple comparisons for L1 and L2 free recall with Tukey's method

	Estimate	SE	z
L1 Free Recall			
Semantic–Mapping	0.16	0.10	1.65
Structural–Mapping	−0.46	0.10	−4.54*
Semantic–Structural	−0.63	0.10	−6.16*
L2 Free Recall			
Semantic–Mapping	−0.22	0.10	−2.22
Structural–Mapping	−0.23	0.10	−2.32
Semantic–Structural	−0.01	0.10	−0.10

Note: SE refers to standard error; * $p < .001$.

present study, however, were the results of the cued recall tests. Both L2-to-L1 cued recall and L1-to-L2 cued recall were higher in the mapping condition than in the semantic condition and in the structural condition, which is consistent with the predictions of the TOPRA model regarding how increases in semantic or structural processing can decrease processing resources available for the mapping component of vocabulary learning. This pattern of results in cued recall can be seen clearly in Figure 3.

DISCUSSION

The main findings of this study are that both semantic and structural tasks substantially decreased performance on the mapping component of L2 vocabulary learning. From a theoretical perspective, these findings are consistent with the predictions of the TOPRA model and provide convincing new support for the model regarding this previously untested component of L2 vocabulary learning. As such, the findings also confirm each of our three hypotheses for the study. From a pedagogical standpoint, the findings point to the importance of allowing learners to process vocabulary as input without requiring them to perform tasks that invoke extraneous types of processing that ultimately exhaust processing resources in directions other than the learning task at hand. More specific to the mapping condition highlighted in the present study, language instructors and developers of instructional materials should not expect semantic or structurally

TABLE 5.   Results of logit mixed-effects model for L2-to-L1 and L1-to-L2 cued recall

Fixed Predictor	Estimate	SE	z
L2-to-L1 Cued Recall			
Intercept	1.53	0.23	6.52*
Condition (Semantic vs. Mapping)	−0.85	0.12	−7.11*
Condition (Structural vs. Mapping)	−1.07	0.12	−8.99*
L1-to-L2 Cued Recall			
Intercept	2.59	0.28	9.24*
Condition (Semantic vs. Mapping)	−0.92	0.14	−6.45*
Condition (Structural vs. Mapping)	−0.82	0.14	−5.71*

Note: SE refers to standard error; * $p < .001$.

TABLE 6.   Multiple comparisons for L2-to-L1 and L1-to-L2 cued recall with Tukey's method

	Estimate	SE	z
L2-to-L1 Cued Recall			
Semantic–Mapping	−0.85	0.12	−7.11*
Structural–Mapping	−1.07	0.12	−8.99*
Semantic–Structural	−0.22	0.11	−2.01
L1-to-L2 Cued Recall			
Semantic–Mapping	−0.92	0.14	−6.45*
Structural–Mapping	−0.82	0.14	−5.71*
Semantic–Structural	0.10	0.13	0.80

Note: SE refers to standard error; * $p < .001$.

elaborative tasks to be the most effective means of promoting development of the mapping component of L2 vocabulary knowledge. To the contrary, they should expect such tasks to *decrease* performance in this area, at least regarding the two tasks examined to date. The following sections discuss the theoretical and pedagogical implications of the study in further detail.

THEORETICAL IMPLICATIONS

As mentioned previously, the present findings provide evidence in support of the TOPRA model from a new angle. Although the TOPRA model has predicted since its inception that the mapping component of vocabulary learning is dissociable from the formal and semantic components and that increases in semantic or structural processing can decrease processing and learning for the mapping component, these predictions had not been empirically tested so directly prior to the present study. From a cognitive
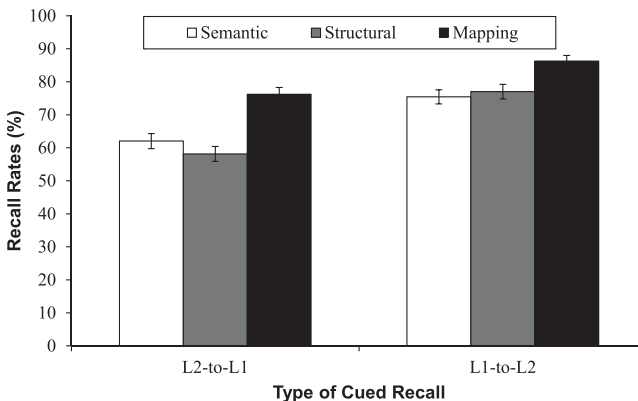

FIGURE 3.   Results of L2-to-L1 and L1-to-L2 cued recall indicating better learning performance in the mapping condition over the other two conditions.
     Note: Error bars are standard errors.

standpoint, the predictions of the model in this regard emphasize specificity in type of processing and allocation of limited processing resources. Because the mapping component of vocabulary learning is—however largely, completely, or incompletely—dissociable from the semantic and structural components of vocabulary learning, the model predicts that as either semantic or structural processing increases, the inner bars in the model must widen the space allocated for those types of processing and necessarily decrease space for other types of processing, in this case, space for the mapping component (see Figure 4 for a visual depiction of these processes). As space for the mapping component decreases, so does learning performance regarding the mapping component, which is precisely what the results for cued recall in the study bore out.

While free recall was not the primary focus of this study, we note that the results for free recall in L1 revealed a typical LOP effect (higher means for the semantic over the structural condition), which helps to confirm that the pleasantness ratings task in the study was indeed invoking the type of processing intended and doing so to a sufficient degree. We also believe that the lower means for L1 free recall in the structural condition as compared to the mapping condition make sense given that in the mapping condition learners were focusing on the mapping between known L2 words and different L1 translations of them, which could lead to better memory for the L1 words in question. In the mapping condition, they could do so without being distracted by the structural task, which promoted a type of processing that was inconsistent with the mapping task being performed. This lack of

A The TOPRA model for increased semantic processing
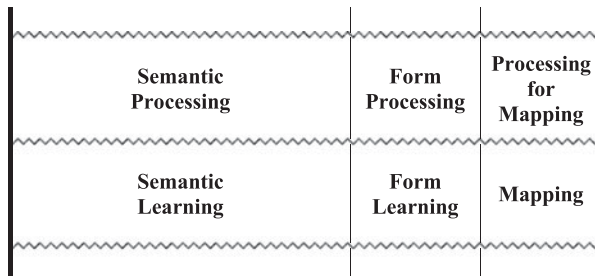
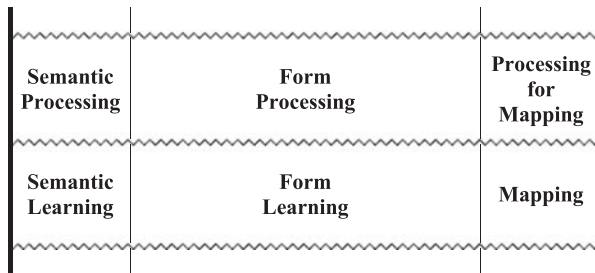B The TOPRA model for increased structural (form) processing

FIGURE 4. The TOPRA model three components of vocabulary learning in contexts of increased semantic processing and increased structural (form) processing.

distraction may have resulted in better memory, as reflected in the higher free recall means in the mapping condition as compared to the structural condition.

How completely or incompletely dissociable is the mapping component of L2 vocabulary learning from the semantic and structural components? This is an important question that needs to be addressed considering the larger body of research on different types of tasks, different types of processing, and learning outcomes. The combined findings of previous studies on the predictions of the TOPRA model regarding different semantic and structural tasks (as surveyed in the review of previous research) strongly support a high degree of dissociability between semantic and structural components of L2 vocabulary learning. If it were not the case that these two types of processing are largely dissociable, one would not expect findings such as those of Barcroft (2002) showing an inverse LOP effect for L2 free recall or those of Wong and Pyun (2012) showing strong negative effects for a semantically oriented sentence-writing task on learning novel word forms (and even more pronouncedly so for a language with forms that are more novel to the L2 learners [Korean as compared to French for English speakers]). The present findings provide initial but strong support in favor of a high degree of dissociability between the semantic and mapping components and between the structural and mapping components of L2 vocabulary learning. If it were not the case that these types of processing were largely dissociable, one would not expect the semantic and structurally oriented tasks to produce such a pronounced negative effect on mapping performance as the one observed in the present study.

Are the semantic, structural, and mapping components of L2 vocabulary wholly and completely dissociable? As phrased, this question remains somewhat elusive based on behavioral data currently available, but the evidence to date suggests that these three types of processing are indeed at least very largely dissociable and that in no way should one expect increasing one of these types of processing to produce learning outcomes that are tied to another one of the three. Even if future research provides evidence indicating a blurring of lines between, for example, the semantic and mapping components (which are dissociated in the TOPRA model), evidence of such blurring would not negate the present data indicating that the two components are, to a substantial if not large or complete degree, dissociable.

At present, the bulk of the evidence suggests that distinguishing among the semantic, structural, and mapping components of L2 vocabulary learning is critical to making accurate predictions about the relationship between different types of tasks, the types of processing they generate, and different possible learning outcomes related to intentional L2 vocabulary learning. The situation is more complex when it comes to incidental L2 vocabulary learning as the impact of tasks in this context potentially can draw learners' attention to words that they might otherwise ignore (see the attention-drawing hypothesis, Barcroft, 2009), but there is evidence to suggest that TOPRA may be applicable in different ways when it comes to incidental L2 vocabulary learning as well (Barcroft, 2009; Kida, 2010).

Future research can continue to address several questions in this area. To begin, the present study examined only one semantic task (pleasantness rating) and one structural task (letter counting). While these tasks have a long history in research on LOP, it should be possible to test other tasks of varying degrees of intensity to examine their effects on

the mapping component of L2 vocabulary learning. Instead of letter counting, for example, one might examine the effect of requiring participants to recode the forms of English words in participants' L1 if the L2 and L1 have different scripts such as English and Japanese or to perform some type of rhyming task (see, e.g., Morris et al., 1977) instead of the letter-counting task. The task was used by Kida (2010), who had Japanese-speaking L2 English learners write down possible pronunciations of English target words by using the Japanese script kana. The English word *tusk*, for example, might be recoded (transcribed) as タスク using Japanese kana sequences. This recoding task results in more intense structurally oriented processing and, therefore, may result in an even greater negative effect on the mapping of secondary meanings of homographs. Next, and importantly, if future studies can identify tasks that focus on the mapping component of L2 vocabulary learning, these tasks may be expected to produce positive learning outcomes for the mapping component tested in the present study and at the same be expected to produce potential negative effects on alternative vocabulary learning tasks that depend on the semantic or structural component. Finally, in our view, homograph learning offers several benefits when it comes to research methodology. As we have argued and attempted to clarify, success in homograph learning as carried out in the present study isolates the mapping component of L2 vocabulary learning because no novel word forms or meanings need to be learned in homograph learning of this nature. Researchers may take advantage of this task to examine a variety of other issues related to the mapping component of vocabulary learning in future studies.

PEDAGOGICAL IMPLICATIONS

From an instructional standpoint, the most straightforward implication of this study is that language instructors, course coordinators, language program directors, and developers of instructional materials should not expect semantic and structural tasks to facilitate learning the mapping component of L2 vocabulary. Instead, what should be more effective is to allow learners opportunities to process the target words without imposing tasks that involve types of processing inconsistent with what is needed for the particular learning goal in question. It certainly may be possible to develop tasks that would promote mapping learning of L2 homographs, but more research is needed to determine precisely which tasks work well. Until then, one at least can surmise that is better to require no concurrent task than to require a semantic or structural task if the mapping component of L2 vocabulary learning is the goal.

We acknowledge that the types of tasks used in the present study are not tasks that are commonly used in the L2 classroom. They were selected for experimental expediency and comparative purposes given their prominence in previous research within the LOP framework and beyond. As such, they were selected with a focus on the larger issue of the effects of semantic and structurally oriented processing and the extent to which these types of processing might be dissociable from the mapping component of L2 vocabulary from a TOPRA-oriented perspective. Future studies may, therefore, provide more directly applicable pedagogical implications in this area by assessing the effects of semantic and structural tasks that are more representative of classroom practices. Also, the present study included only tests that were administered immediately after the study phase. Future studies can include delayed posttests, such as with delays of two weeks or

more, to determine whether the effects observed in the present study change as a function of delayed time.

From a larger perspective, the combined findings of this study and previous studies on the TOPRA model suggest that making decisions about effective vocabulary instruction does not have to be a guessing game. We can investigate the relative effectiveness of different types of tasks when it comes to different aspects of vocabulary learning. To date, research findings in this area have unveiled numerous findings that allow instructors to make more informed decisions when it comes to L2 vocabulary instruction. Semantic tasks such as sentence writing or answering questions about word meaning should not be expected to facilitate word form learning (Barcroft, 2003, 2004; Wong & Pyun, 2012), for example. As another example, structural processing tasks such as recoding L2 words into known writing scripts may be beneficial in some contexts (Kida, 2010). Findings such as these have direct implications regarding what is and is not effective for vocabulary instruction, as do the present findings regarding the negative effects of at least one semantic and one structural task on the mapping component of L2 vocabulary learning.

NOTES

[1]Regarding etymology, *homonym* refers to having a same name (from Greek, *homonumos*) and *homograph* (also from Greek) to having the same writing. For some target words in the present study, such as *dish*, *floor*, and *dog*, a polysemous relationship can also be identified (although perhaps not readily, e.g., the case of *dish* in the sense of "serving up" information or gossip), making them "polysemous homonyms" and "polysemous homographs." Although some might argue against the use of these terms when it comes to technical classification, for the purposes of the present study, the key point is that the alternate meanings of these words were substantially different (much more so, e.g., than in the clearly related case of the English word *talk* as a verb and *talk* as a noun) and had translations in L1 that were already known based on years of previous experience in L1.

[2]Mapping secondary meanings of L2 homographs through L1 translation in no way guarantees that a learner establishes fully appropriate L2-specific knowledge of the secondary meaning of the homograph in question because complete one-to-one translation correspondence *across* languages should not be expected. Nonetheless, what matters for the purposes of the present study is that learning secondary meanings of L2 homographs through L1 translation depends on the mapping component (even though the secondary meaning in question does not include all the L2-specific semantic space, collocational properties, and usage patterns that one acquires over time through usage-based learning outside of the context of translation-based learning alone).

[3]An alternate case is when learners must encode and retain a *new* L2 word meaning as well as a new form-meaning mapping, such as when a Japanese speaking undergraduate student studying linguistics first learns the English word *stop* (閉鎖音) in a phonetics class after having previously learned the more common meaning of this word (停止). In this case, *new* meaning is being learned because the concept in question did not previously exist in the learner's mental lexicon. In contrast to cases such as the example of *stop* provided here, the learning task in the present study focused on mapping *known* L2 forms to *known* meanings previously acquired in L1.

[4]When knowledge of L2 word form is partial, a learner could respond correctly in L2-to-L1 cued recall but could respond only partially (at best) in the case in L1-to-L2 cued recall, the former situation sometimes being called receptive vocabulary learning and the latter sometimes being called productive vocabulary learning, but from another perspective, the productive-receptive distinction is actually only a reflection of amount of word form knowledge and how it is accessed (by either providing the L2 word form and allowing the learner to map partial word form knowledge onto it or by not providing the L2 word form and requiring the learner to produce it). Note, however, that previous research (as reviewed in the preceding text) indicates that increased semantic processing (invoked by semantic tasks such as sentence writing and addressing questions about word meaning)

has resulted in decreases in both L1-to-L2 *and* L2-to-L1 cued recall, which is inconsistent with the idea of increased semantic processing somehow encouraging "receptive" vocabulary learning more so than "productive" vocabulary learning.

[5]The focus of our review of research literature here is intentional L2 vocabulary learning. Other studies (Barcroft, 2009; Burfoot, 2010; Kida, 2010) have attempted to test the applicability of TOPRA predictions regarding the effects of semantic and structural tasks in more incidentally oriented contexts of L2 vocabulary learning. Like the studies on TOPRA regarding intentional L2 vocabulary learning, none of these studies revealed any positive effect for any of the semantically oriented tasks that they assessed. Kida (2010), however, found that having Japanese-speaking L2 English learners recode English word forms by writing a "possible" pronunciation of each word in the kana writing system (the learners generated the "possible" pronunciations on their own) positively affected incidentally oriented L2 word learning.

[6]The original participant pool for the study included 124 Japanese-speaking L2 English learners. Four participants were excluded because they did not complete all the experimental procedures. Furthermore, data provided by eight participants were excluded because they did not perform the experimental tasks correctly or as intended (e.g., circling the same number for every target word).

[7]Lexical characteristics we controlled among these three sets were (a) the number of letters and the objective frequency of the English words (based on CELEX, analyzed by the N-Watch software, Davis, 2005) and (b) the frequency and the number of moras of Japanese secondary translations (based on *Shonagon* database of Balanced Corpus of Contemporary Written Japanese available at http://www.kotonoha.gr.jp/shonagon/). Moras are units of the sound system used in Japanese, somewhat comparable to syllables in English. They depict the length of the sound of each word. For example, 足 is pronounced *a shi* and has two moras whereas 支払う is pronounced *shi ha ra u* and has four moras. One-way ANOVAs revealed no significant difference among the three sets regarding the (a) number of letters, $F(2, 21) = 0.09, p = .92$, (b) the objective frequency of English words, $F(2, 21) = 0.05, p = .96$, (c) the frequency of Japanese translations, $F(2, 21) = 0.03, p = .97$, and (d) the number of moras, $F(2, 21) = 0.15, p = .86$.

REFERENCES

Barcroft, J. (2002). Semantic and structural elaboration in L2 lexical acquisition. *Language Learning*, *52*, 323–363. doi: 10.1111/0023-8333.00186

Barcroft, J. (2003). Effects of questions about word meaning during L2 lexical learning. *The Modern Language Journal*, *87*, 546–561. doi: 10.1111/1540-4781.00207

Barcroft, J. (2004). Effects of sentence writing in L2 lexical acquisition. *Second Language Research*, *20*, 303–334. doi: 10.1191/0267658304sr233oa

Barcroft, J. (2007). Effects of opportunities for word retrieval during second language vocabulary learning. *Language Learning*, *57*, 35–56. doi: 10.1111/j.1467-9922.2007.00398.x

Barcroft, J. (2009). Effects of synonym generation on incidental and intentional L2 vocabulary learning during reading. *TESOL Quarterly*, *43*, 79–103. doi: 10.1002/j.1545-7249.2009.tb00228.x

Barcroft, J. (2015). *Lexical input processing and vocabulary learning*. Amsterdam, The Netherlands: John Benjamins.

Bartoñ, K. (2015). MuMIn: Multi-Model Inference. R package version 1.13.4. Retrieved from http://CRAN.R-project.org/package=MuMIn

Bates, D., Maechler, M., Bolker, B. M., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. doi: 10.18637/jss.v067.i01

Bogaards, P. (2001). Lexical units and the learning of foreign language vocabulary. *Studies in Second Language Acquisition*, *23*, 321–343. doi: 10.1017/S0272263101003011

Burfoot, S. (2010). *Effects of synonym generation on incidental and intentional L2 vocabulary learning during reading: A replication of Barcroft (2009)*. Retrieved from http://www.academia.edu/669738/The_Effects_of_Synonym_Generation_on_L2_Vocabulary_Recall

Comesaña, M., Perea, M., Piñeiro, A., & Fraga, I. (2009). Vocabulary teaching strategies and conceptual representations of words in L2 in children: Evidence with novice learners. *Journal of Experimental Child Psychology*, *104*, 22–23. doi: 10.1016/j.jecp.2008.10.004

Coomber, J. E., Ramstad, D. A., & Sheets, D. R. (1986). Elaboration in vocabulary learning: A comparison of three rehearsal methods. *Research in the Teaching of English*, *20*, 281–293.

Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Behavior*, *11*, 671–684. doi: 10.1016/S0022-5371(72)80001-X

Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, *104*, 268–294. doi: 10.1037/0096-3445.104.3.268

Crossley, S., Salsbury, T., & McNamara, D. (2010). The development of polysemy and frequency use in English second language speakers. *Language Learning*, *60*, 573–605. doi: 10.1111/j.1467-9922.2010.00568.x

Davis, C. J. (2005). N-watch: A program for deriving neighborhood size and other psycholinguistic statistics. *Behavior Research Methods*, *37*, 65–70. doi: 10.3758/BF03206399

De Groot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, *50*, 1–56. doi: 10.1111/0023-8333.00110

Durkin, K., & Manning, J. (1989). Polysemy and the subjective lexicon: Semantic relatedness and the salience of intraword senses. *Journal of Psycholinguistic Research*, *18*, 577–612. doi: 10.1007/BF01067161

Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, *43*, 559–617. doi: 10.1111/j.1467-1770.1993.tb00627.x

Finkbeiner, M., Forster, K., Nicol, J., & Nakamura, K. (2004). The role of polysemy in masked semantic and translation priming. *Journal of Memory and Language*, *51*, 1–22. doi: 10.1016/j.jml.2004.01.004

Harrington, M., & Jiang, W. (2013). Focus on the forms: From recognition practice in Chinese vocabulary learning. *Australian Review of Applied Linguistics*, *36*, 132–145.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, *50*, 346–363. doi: 10.1002/bimj.200810425

Hulstijn, J. H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. J. Arnaud, & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 113–125). London, UK: Macmillan.

Hyde, T. S., & Jenkins, J. J. (1969). The differential effects of incidental tasks on the organization of recall of a list of highly associated words. *Journal of Experimental Psychology*, *82*, 472–481. doi: 10.1037/h0028372

Jiang, N. (2000). Lexical representation and development in a second language. *Applied Linguistics*, *21*, 47–77. doi: 10.1093/applin/21.1.47

Kida, S. (2010). The role of processing-resource allocation in incidental L2 vocabulary learning through reading. *Annual Review of English Language Education in Japan*, *21*, 171–180.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. New York, NY: Longman.

Laufer, B. (1997). What's in a word that makes it hard or easy: Some intralexical factors that affect the learning of words. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 140–155). Cambridge: Cambridge University Press.

Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning*, *65*, 185–207. doi: 10.1111/lang.12117

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. doi: 10.1111/1467-9922.00032

McNamara, D. S., & Healy, A. F. (1995). A generation advantage for multiplication skill training and nonword vocabulary acquisition. In A. F. Healy & L. E. Bourne Jr. (Eds.), *Learning and memory of knowledge and skills: Durability and specificity* (pp. 132–169). Thousand Oaks, CA: Sage.

Morimoto, S., & Loewen, S. (2007). A comparison of the effects of image-schema-based instruction and translation-based instruction on the acquisition of L2 polysemous words. *Language Teaching Research*, *11*, 347–372. doi: 10.1177/1362168807081181

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Levels of processing versus transfer appropriate processing. *Journal of Verbal Learning and Behavior*, *16*, 519–533. doi: 10.1016/S0022-5371 (77)80016-9

Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133–142. doi: 10.1111/j.2041-210x.2012.00261.x

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nation, P. (2007). Fundamental issues in modelling and assessing vocabulary knowledge. H. Daller, J. Milton, & J. T. Daller (Eds.), *Modelling and assessing vocabulary knowledge* (pp. 35–43). Cambridge: Cambridge University Press.

Pressley, M., Levin, J. R., & Miller, G. E. (1982a). The keyword method compared to alternative vocabulary-learning strategies. *Contemporary Educational Psychology*, 7, 50–60. doi: 10.1016/0361-476X(82)90007-8

Pressley, M., Levin, J. R., Kuiper, N. A., Bryant, S. L., & Michener, S. (1982b). Mnemonic versus non-mnemonic vocabulary-learning strategies: Additional comparisons. *Journal of Educational Psychology*, 74, 693–707. doi: 10.1037/0022-0663.74.5.693

Prince, P. (1996). Second language vocabulary learning: The role of context versus translation as a function of proficiency. *The Modern Language Journal*, 80, 478–493. doi: 10.1111/j.1540-4781.1996.tb05468.x

R Development Core Team. (2013). R: A language and environment for statistical computing [Computer software]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Roediger, H. L., & Guynn, M. J. (1996). Retrieval processes. In E. L. Bjork & R. A. Bjork (Eds.), *Memory* (pp. 197–236). San Diego, CA: Academic Press.

Royer, J. M. (1973). Memory effects for test-like events during acquisition of foreign language vocabulary. *Psychological Reports*, 32, 195–198. doi: 10.2466/pr0.1973.32.1.195

Ryan, A. (1997). Learning the orthographical form of L2 vocabulary: A receptive and a productive process. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition, and pedagogy* (pp. 181–198). Cambridge: Cambridge University Press.

San Mateo Valdehíta, A. (2013). El efecto de tres actividades centradas en las formas (Focus on forms, fonfs): la selección de definiciones, la selección de ejemplos y la escritura de oraciones, en el aprendizaje de vocabulario de segundas lenguas [The effect of three form-focused activities (Focus on forms, fonfs): selecting definitions, selecting examples, and sentence writing, on second language vocabulary learning]. *Revista Electrónica de Lingüística Aplicada*, 12, 17–36.

Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, 48, 281–317. doi: 10.1111/1467-9922.00042

Schwartz, B. D. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition*, 15, 147–163. doi: 10.1017/S0272263100011931

Stein, B. S. (1978). Depth of processing reexamined: The effects of the precision of encoding and test appropriateness. *Journal of Verbal Learning and Verbal Behavior*, 17, 165–174. doi: 10.1016/S0022-5371(78)90128-7

Storkel, H. L., & Maekawa, J. (2005). A comparison of homonym and novel word learning: The role of phonotactic probability and word frequency. *Journal of Child Language*, 32, 827–853. doi: 10.1017/S0305000905007099

Tinkham, T. (1993). The effect of semantic clustering on the learning of second language vocabulary. *System*, 21, 371–380. doi: 10.1016/0346-251X(93)90027-E

Verspoor, M., & Lowie, W. (2003). Making sense of polysemous words. *Language Learning*, 53, 547–586. doi: 10.1111/1467-9922.00234

Wang, A., Thomas, M. H., & Quellette, J. A. (1992). They keyword mnemonic and retention of second language vocabulary. *Journal of Educational Psychology*, 84, 520–528. doi: 10.1037/0022-0663.84.4.520

Waring, R. (1997). The negative effects of learning words in semantic sets: A replication. *System*, 25, 261–274. doi: 10.1016/S0346-251X(97)00013-4

Waring, R., & Takaki, M. (2003). At what rate do learners learn and retain new vocabulary from reading a graded reader? *Reading in a Foreign Language*, 15, 130–163.

West, M. (1953). *A general service list of English words*. London, UK: Longman.

Wong, W., & Pyun, D. O. (2012). The effects of sentence writing on second language French and Korean lexical retention. *The Canadian Modern Language Review*, 68, 164–189. doi: 10.3138/cmlr.68.2.164

APPENDIX

THE EXPERIMENTAL MATERIALS USED IN THE STUDY

	1st Meaning	Translation	2nd Meaning	Translation
head	on neck	頭	head north	進む
bowl	container	おわん	throw a ball	転がす
shell	seashell	貝殻	shell the enemy	砲撃する
cow	animal	牛	cow into submission	脅す
trip	vacation	旅行	to stumble	つまずく
coast	coast guard	海岸	coast to victory	楽勝する
floor	ground floor	床	floor someone	閉口させる
dog	not cat	犬	to follow	尾行する
bolt	screw	ねじ	run away	逃亡する
plug	electricity	コンセント	to advertise	宣伝する
dish	container	皿	to gossip	噂する
man	male	男	man a ship	配置する
sock	clothing	靴下	to hit	強打する
stump	tree	切り株	to puzzle	困らせる
stem	plant	茎	stem the tide	止める
dress	clothes	服装	hunted animal	下処理する
foot	not hand	足	foot the bill	支払う
court	judge	法廷	court the princess	言い寄る
snow	not rain	雪	to cheat	騙す
line	draw a line	線	line the wall	覆う
ice	frozen	氷	to kill	殺す
rock	stone	岩	rock a baby	揺らす
case	box	箱	case the area	下調べする