

The Organization of the Senses of Polysemy in Japanese EFL Learners' Mental Lexicon

Yuko Hoshino¹, Haruka Shimizu²

¹Department of Teacher Education, Shumei University, Yachiyo, Japan

²Department of Education, Tohoku Fukushi University, Sendai, Japan

Email: hoshino@mailg.shumei-u.ac.jp

How to cite this paper: Hoshino, Y., & Shimizu, H. (2018). The Organization of the Senses of Polysemy in Japanese EFL Learners' Mental Lexicon. *Creative Education*, 9, 353-367.

<https://doi.org/10.4236/ce.2018.93025>

Received: January 25, 2018

Accepted: March 9, 2018

Published: March 12, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution-NonCommercial International License (CC BY-NC 4.0).

<http://creativecommons.org/licenses/by-nc/4.0/>



Open Access

Abstract

The structure of the mental lexicon has been widely researched, but few studies focusing on polysemy have been conducted, even in an L1 (a learner's first language) context, and almost no research has been conducted in an L2 (a learner's second language) context. The current study aims to scrutinize how the different vocabulary size groups of Japanese EFL learners classify the various senses of basic polysemous words and to compare their categorization with sense classification based on a linguistic dictionary. The results indicated that those with an estimated vocabulary size consisting of 5500 words or over (hereafter, the upper group) classified senses in a more similar way to the dictionary than those whose vocabulary size comprised 5000 or fewer words (hereafter, the lower group). This was despite the fact that they both understood the target context well and the number of categorizations created was almost the same between the two groups. Moreover, the upper group tended to classify the senses of polysemy in a more similar way to the other participants in the same vocabulary group, while the lower group's categorization was more divergent and less well-organized. These findings indicated that even though the participants understood each sense, the difference appeared in the ability to categorize the senses, suggesting that grouping the senses needs knowledge different from simply understanding them, and this knowledge is not yet fully developed even for basic words.

Keywords

EFL, Polysemy, Mental Lexicon, Vocabulary Acquisition, Senses

1. Introduction

Language learners have been said to have lexical networks in their mental lexicon

and words in the mental lexicon are linked with each other. Research by [Collins and Quillian \(1969\)](#) was probably the first to advocate a network model; they proposed the Hierarchical Network Model. Later, [Collins and Loftus \(1975\)](#) suggested the Spreading Activation Model, in which words are connected in the mental lexicon when they have semantic relationships. Based on these models, it is now considered a matter of course that words have connections with other words in the mental lexicon. There has been lots of research done on these lexical networks over the last 50 years (e.g., [Nissen & Henriksen, 2006](#); [Wolter, 2001, 2006](#)).

However, the lexical network is constructed not only between words but also within a word itself. That is, many words, especially high frequency words, have more than one meaning, and each meaning has certain relationships with the other meanings within the word. There are two categorizations to describe words with multiple meanings: homonymy and polysemy. Homonymy is defined as a word having two or more meanings with no semantic relationship (e.g., *bank* as an organization that provides various financial services vs. the side of a river), so each meaning of a homonym accidentally shares the spelling and sound with the others. Furthermore, the meanings of homonyms usually do not have the same origin. On the other hand, polysemous words have multiple meanings which have semantic relationships with each other (e.g., *have* has several meanings such as “to own,” “to experience,” “to eat”). Since homonymy and polysemy have different characteristics, some researchers use the word “sense” to refer to the meanings of polysemy, but not “meaning” (e.g., [Foraker & Murphy, 2012](#)). In this paper, we use “sense” to indicate the meaning of a polysemous word in order to avoid confusion.

Just as there is a clear difference between homonymy and polysemy, as shown above, the structures of these two types of word are thought to be different in our mental lexicon. For instance, L1 users responded faster to polysemy than homonymy in a lexical decision task ([Rodd, Gaskell, & Marslen-Wilson, 2002](#)). Furthermore, in a lexical decision task conducted in the participants’ L1, they also responded faster to polysemous words than to words with a single meaning ([Klepousniotou, 2002](#)). From these results, we can hypothesize that the structure of polysemy in the L1 mental lexicon is somewhat different from both homonymy and words with only one meaning.

Then, how do human beings perceive the senses of polysemy? Based on the idea of cognitive linguistics, words with multiple senses have a core sense, and other senses, such as metonymy, metaphor, and synecdoche, are extended from the core sense. [Langacker \(1999\)](#) presented an example of the network structure of polysemy as shown in [Figure 1](#). In this figure, A, B, C, and D represent each sense of a polysemous word, with A as the core sense. A’ and C’ are schematic extended senses, and A₁, A₂, and A₃ are the example usages of sense A. Arrows are a constructed network between each node, and arrows with broken lines indicate that the connections between nodes are not fully established, compared to

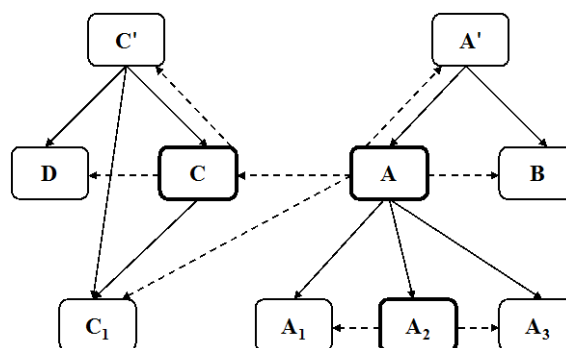


Figure 1. An example of polysemic construction based on network model (Modified from Langacker, 1999: p. 103).

the arrows with solid lines. To be more specific, the mental lexicon described in **Figure 1** shows that the connections between A and B, or A and C, are to some degree constructed, but the link between A and D has not been established, since there is no arrow. Theoretically, this network seems to work well, but it has not yet been confirmed whether human beings have such a network of senses. For example, even if A, B, C, and D are senses described in dictionaries, learners might classify senses more roughly or more precisely than dictionaries. In other words, it is possible for learners to interpret A and B as the same sense and put A and B in the same node, or alternatively, learners may divide A into more detailed categories.

In L1 vocabulary acquisition, Aitchison (2003) suggested that there were three levels: labeling, packaging (categorization), and network building. Children first link the referent and the word (e.g., knowing that *duck* is a kind of yellow bird swimming in a pond), and then secondly, they know many types of that word (for example, a swimming duck, walking duck, a big duck, and a little duck can all be defined by the word *duck*). After that, they link the word with other words (such as linking *duck* with other birds like *hen*, *swan*, and *goose*). Although Aitchison's suggestion is not only targeted at L2 learners and not only focused on polysemy, it is assumed that L2 learners first recognize the new polysemic word and understand some of its senses, and that they then differentiate one sense from others (constructing A, B, C, D in **Figure 1**). After that, they connect the link between each sense (shown by the arrows in **Figure 1**). However, **Figure 1** is a theoretical model and there is no empirical proof about how EFL learners recognize different senses in one word. In contrast to native English speakers, EFL learners' mental lexicon is affected by various learners' factors such as English proficiency, vocabulary knowledge or their L1 (e.g., Jiang, 2000).

Other than learners' factors, one more factor to consider in constructing these nodes of polysemy in the mental lexicon is how much each sense described in a dictionary overlaps with the others. In this case, how much semantic overlap is required to be recognized as one sense? Some researchers have already investigated this issue. For example, Brown (2008) presented the target words with one of the senses among the same senses (e.g., *cleaned the shirt* and *cleaned the cup*),

closely related senses (e.g., *broke the glass* and *broke the video*), distantly related senses (e.g., *ran the track* and *ran the shop*), and different senses (e.g., *banked the plane* and *banked the money*). The participants judged whether each phrase made sense or not as quickly and accurately as possible. The results showed that when the senses were more closely related, the participants responded faster and more accurately. This result indicated that the similar senses were connected more closely in the lexical network. Klepousniotou, Titone, and Romero (2008) also showed that the context effect lessened when senses were highly overlapped with each other (e.g., *lamb* referring to the animal or to the meat of that animal), compared to when senses only moderately overlapped (e.g., *beam* referring to a piece of wood or to a ray of light). This means that whatever the context, senses were recognized as making one sense when they were highly overlapped. These two studies indicated that highly overlapping senses were more easily understood than moderately- or low-overlapping senses, and highly overlapping senses are tied more strongly to each other than other senses. The important implication from these two studies was that it is not only the semantic overlap between words but also the semantic overlap between the senses that affects lexical networks. Unfortunately, no such studies using polysemy have been conducted on L2 so far. Compared to an L1 user's mental lexicon, that of an L2 is affected by a number of other factors. For example, the structure of the mental lexicon itself is known to differentiate by the participants' L2 proficiency (e.g., Zareva, 2007). However, little is known about the construction of polysemy in the L2 mental lexicon.

Hence, this study investigates how Japanese EFL learners regard various senses as the same sense, that is, focusing on the packaging stage in Aitchison (2003), and how this is differentiated by their vocabulary size. In order to investigate these issues, the current study examined similarities between EFL learners' cognitive categorizations and a dictionary's linguistic categorization. When words have multiple senses, linguistic specialists make the categorizations described in dictionaries. But no research has been conducted into whether EFL learners categorize the senses of words in a similar way to the dictionary categorization. Some L1 studies noted that the definitions in dictionaries do not always have strong relationships with learners' mental lexicons. For example, Azuma and van Orden (1997) found that the number of senses described in dictionaries did not influence the participants' lexical access. That is, even if a polysemous word has more multiple senses than another polysemous word, the times required for lexical access between them were not different. Lin and Ahrens (2010) also showed that the number of senses that their participants produced differed from the number of senses described in dictionaries. These findings indicate that the number of senses in dictionaries does not seem to directly reflect the mental lexicon. However, the relationship between dictionaries and the mental lexicon in L2 has not been investigated. Moreover, many of the studies concerning words with multiple meanings adopted only two senses (e.g., Frazier & Rayner,

1990; Klepousniotou, Titone, & Romero, 2008). This is especially true when studies compared polysemy and homonymy which usually have two distant meanings. However, polysemous words, especially basic words, have many senses, and descriptions of only the top 500 high frequency words occupy more than 50% of a learners' dictionary containing 100,000 words (Tanaka, 2016). Although using only two senses for each word can increase the possible number of words to investigate in one study, examining as many senses as possible in one word is also beneficial for clarifying the structure of polysemy in the mental lexicon.

The purpose of the current study was to investigate how Japanese EFL learners classify sentences containing several senses of a basic polysemous word. We prepared three or more senses for each target word and focused on the packaging stage in Aitchison (2003). The categorizations made by two groups of Japanese EFL learners with different vocabulary sizes and that of the dictionary were compared in order to find the similarities and differences between the dictionary and the learners' mental lexicons. Two research questions (RQs) were set to investigate these issues:

RQ1: What kinds of differences were found between the sense categorizations made by the dictionary and those made by Japanese EFL learners?

RQ2: Are there any differences in sense categorization in relation to the English vocabulary level of Japanese EFL learners?

2. Method

2.1. Participants

A total of 73 university students voluntarily participated in this study. All of them were Japanese learners of English and belonged to the same university, but their major varied in subjects such as education, social welfare, and psychology. One participant could not take the vocabulary size test, so the data of 72 participants were analyzed. All participants agreed to take part in the experiment.

2.2. Materials

A vocabulary size test and a sense categorization task were adopted. The bilingual version (English to Japanese) of the Nation's Vocabulary Size Test was used to estimate the participants' vocabulary size. Originally, this test was created up to Level 14 (14,000-word level), but the participants in this study were not proficient enough to take all levels, so this study used Levels 1 to 10, which had a total of 100 questions.

In the sense categorization task, six basic English words (i.e., *go*, *look*, *time*, *take*, *get*, *have*) were used to examine participants' lexical networks. According to Hoshino (2016), a study that focused on the senses of polysemy and investigated how many times each sense of polysemy appeared on the Eiken test (a widely used English test in Japan), there were 22 content words within both the top 100 frequency words in the JACET8000 word list (JACET, 2003) and the Dictionary of English Lexical Polysemy (Seto, 2007). We chose 6 words from

those 22 using the following criteria: (a) these words appear with several senses both for upper- and lower-grades on Eiken tests (i.e., from A1 to C1 levels in CEFR: Common European Framework of Reference for Languages), (b) these words have more than two senses, and (c) contexts including target words are not highly biased (that is, whether the contexts give too many contextual cues so learners can guess the sense of the target word). The focus of this study is the packaging stage in Aitchison (2003), so it was necessary that all learners had already passed the labeling stage. Hence, we assumed that participants in the present study were familiar with these words and that the senses of these target words had already been packaged in their mental lexicon at least to some degree. Considering the fatigue effect, we limited the number of target words to six words in the current study because the participants had to read at least nine sentences for each word (see below).

Senses were classified based on the definitions in the Dictionary of English Lexical Polysemy (Seto, 2007). Since this dictionary was compiled from the perspective of cognitive linguistics by Japanese experts and clearly describes a core sense and extended senses separately, we considered that this dictionary's definitions are likely to be more similar to a Japanese learner's mental lexicon than those of other dictionaries. Several sentences were selected from the grammar, vocabulary, and reading sections of Eiken pre-2nd, 3rd, 4th, and 5th grade tests (levels A1 to A2 in CEFR), then two raters independently identified a sense based on the Dictionary of English Lexical Polysemy. The agreement ratio was 72.49% and the remaining disagreement was dissolved through discussion. Afterwards, three sentences per sense were extracted from Eiken tests and some expressions were slightly modified as necessary. The number of senses between target words was not controlled in the current study because the number of senses that original target words have in a dictionary varies according to the word. We used three senses (nine sentences) for *go*, *look*, and *time*, four senses (12 sentences) for *take*, five senses (15 sentences) for *get*, and seven senses (21 sentences) for *have*. In total, there were 75 target sentences. The summary of senses used in this study is shown in Table 1. When target words were verbs, we paid special attention that the objects of the three sentences within a sense did not overlap. This is because the object of the verb seemed to have a strong influence on the participants' decision on the sense of the verb.

2.3. Procedure

The experiment was conducted over two days: a vocabulary test slot and a sense categorization slot. Participants undertook a vocabulary test and a sense categorization task at their own pace. Overall, all participants completed the vocabulary test within 40 minutes and the sense categorization task within 60 minutes. In the vocabulary test slot, the participants chose one Japanese word which matched the English target word from four choices. In the sense categorization slot, before starting the task, the examiner explained the procedure with an

Table 1. Target senses adopted in the current study.

Target	Definition	Example sentence
go	move from one place to another	It's late. I have to be going.
	(of time) pass or elapse	Time goes so fast.
	intend or be likely	I'm going to be late for work.
look	direct one's gaze in a specified direction	He looked at her.
	regard in a specified way	There is another way of looking at this issue.
	have the appearance or give the impression of being	He looks tired.
time	time as allotted, available, or used	You waste too much time.
	an instance of something happening or being done	They heard the gospel for the first time.
	a moment or definite portion of time	The probable meeting time will be 7:00 p.m.
take	carry or bring with one	Let me take your coat.
	possess, buy	He took the Nobel Prize in 2003.
	perform an action	taking a nap
	require of using up	He took 7 years to decide to marry me.
get	receive, hold	They got a tiger by the tail.
	buy	I'll get this one. Put it on my card.
	move or travel	getting home
	change or cause to reach a specified state or condition	I get myself ready for school.
	succeed in attaining, achieving, or experiencing	They're glad to get the opportunity to work here.
have	possess or own things	I had a carrier bag in my pocket.
	used to indicate a particular relationship	My brother has two children.
	possessed by a human (a quality, characteristic, or feature)	The child has dark hair and blue eyes.
	eat or drink	Have another drink.
	experience	Did you have a good time tonight?
	be strongly recommended to do something	You have to come and see us.
	be used as the perfect tense	They have just finished a report.

example, then asked participants to answer example sentences including *make* as an example target word. In the sense categorization task, the participants were given several sentences including the same target word in context and asked to group these sentences into sense categories. They did not know how many sentences belonged to each sense described in the dictionary. Each target word was presented to participants with twelve blank boxes with the instruction to use as many boxes as they wanted when categorizing the sentences. Furthermore, they could add boxes if the printed boxes in the booklet were not sufficient to categorize the sentences. Then, they assessed how well they understood each sentence by use of a 3-point scale (1: I couldn't understand the sentence; 2: I almost understood the sentence; 3: I completely understood the sentence) in order to confirm that the participants understood the target sentences. An example of the

task is shown in the appendix.

2.4. Data Analyses

Three types of data were collected in the current study: (a) vocabulary size, (b) sense categorization, and (c) the degree of understanding each context. The sense categorization data were further divided into two: the number of sense categories the participants made (i.e., the box number used) and how the senses were grouped.

The participants whose vocabulary sizes were 5500 words or over were regarded as the upper group (with a large vocabulary size), and the participants who knew 5000 words or fewer were grouped as the lower group (with a small vocabulary size). By using a *t* test, we investigated whether the number of sense categories and the degree of understanding data differed between the two groups.

Furthermore, we analyzed the sense categorization data by [Morey and Agresti's \(1984\)](#) adjusted Rand Index, which defines two instances of classification agreement. When the two classifications match completely, the adjusted Rand Index becomes 1, and when the classification is completely different, the index is 0. We compared the upper- and the lower-proficiency groups using this index in two ways: (a) to compare the participants' data with the sense classification of the dictionary; and (b) to investigate whether each participant classified senses similarly to the other participants in the same proficiency group. Therefore, after calculating the adjusted Rand Index, we conducted *t* tests to compare the indices of the two proficiency groups. If a participant did not categorize one third or more sentences of a target word, we regarded the target word as missing values and excluded the data from further analysis. Finally, we used Cohen's *d* as the effect size, such as small (0.20), medium (0.50), and large (0.80) ([Mizumoto & Takeuchi, 2008](#)).

3. Results and Discussion

3.1. Vocabulary Size Test

Table 2 shows the result of the vocabulary size, as estimated from the Nation's Vocabulary Size Test (VST). The average vocabulary size of the 72 participants was 5221.13 words. Among them, those who scored 5500 or more were regarded as the upper group ($n = 27$), and those who scored 5000 or less were regarded as the lower group ($n = 26$) for the sake of convenience. A significant difference between the two groups was confirmed: $t(51) = -12.55$ [95% CI: $-13.26, -9.60$], $p = 0.000$, $d = 3.45$.

3.2. Overall Tendency

The mean of the degree of understanding for each group is shown in **Table 3**. The result indicated a ceiling effect: that is, the figures of the mean score plus 1 Standard Deviation (SD) exceed the full score, which is 3. This means that

Table 2. Estimated vocabulary size of upper and lower groups.

	Mean	Standard Deviation	Standard Error
Upper ($n = 27$)	5777.78	225.89	43.47
Lower ($n = 26$)	4634.62	413.71	81.13

Table 3. The degree of understanding target sentences.

Target word	Group	Degree of Understanding	Standard Deviation
go	Upper	2.64	0.66
	Lower	2.55	0.71
look	Upper	2.65	0.62
	Lower	2.50	0.73
time	Upper	2.79	0.47
	Lower	2.63	0.66
take	Upper	2.60	0.66
	Lower	2.39	0.79
get	Upper	2.78	0.52
	Lower	2.49	0.75
have	Upper	2.68	0.65
	Lower	2.54	0.74

participants understood the target sentences enough to conduct the sense categorization task.

Table 4 shows the number of categories participants created in the sense categorization task. Overall, the participants created more categories than are found in the dictionary. For example, the dictionary defined three senses each for *go*, *look*, and *time* (see n of senses in **Table 4**), but the participants produced an average of over four senses for the same words. In addition, the fewer senses a target word had, the fewer categories participants created. Exceptionally, for *take*, which had four prepared senses from the dictionary's definition, participants produced more than seven categories. This is even larger than the classification of *get*, with five senses given in the dictionary. In terms of the number of groups the participants classified, there were no statistical differences between the upper and lower groups in any of the target words, as summarized in **Table 4**. Therefore, there was no difference in the quantity of classifications between the groups.

3.3. Comparison of the Upper and Lower Groups

Next, Morey and Agresti's (1984) adjusted Rand Index was calculated in order to focus on the quality of the sense categorization between the two groups. As explained in Section 2.4, students who did not categorize more than one third of the senses were excluded from the analysis, so the number of the participants

Table 4. The number of categories made by upper and lower groups.

Target Word	<i>n</i> of senses	Group	<i>M</i>	<i>SD</i>	<i>SE</i>	<i>p</i>	Effect size (<i>d</i>) [95%CI]
go	3	Upper	4.22	0.89	0.17	0.921	0.03 [-0.63, 0.57]
		Lower	4.19	1.27	0.25		
look	3	Upper	4.19	1.14	0.22	0.830	0.06 [-0.72, 0.58]
		Lower	4.12	1.21	0.24		
time	3	Upper	4.04	1.06	0.20	0.924	0.03 [-0.81, 0.74]
		Lower	4.00	1.70	0.33		
take	4	Upper	7.44	2.06	0.40	0.523	0.17 [-1.67, 0.86]
		Lower	7.04	2.52	0.49		
get	5	Upper	6.70	1.86	0.36	0.755	0.08 [-1.08, 1.45]
		Lower	6.89	2.67	0.52		
have	7	Upper	8.15	2.80	0.54	0.819	0.06 [-1.23, 1.55]
		Lower	8.31	2.20	0.43		

slightly differed between the target words.

Table 5 displays a comparison between each participant's categorization and those of the dictionary. As shown by the adjusted Rand Index in **Table 5**, the upper group always had a higher index, meaning that the upper group classified word senses more similarly to the dictionary categorization than the lower group did. The effect size (*d*) of four words was over 0.50, which indicated that there were medium or large effects. It can therefore safely be said that the categorization of polysemy of students with a large vocabulary size was more like the dictionary's classification than that of students with a small vocabulary size.

Table 6 shows to what extent each individual participant's categorization is similar to the other participants in the same group. As can be seen, the effect of vocabulary size was significant and all target words had an effect size of over 0.50, which indicated that the upper students classified senses more similar to the other participants in the same group than in the lower group. In the lower group, each individual classified senses uniquely, and the classifications were not consistent within the group. Upon integrating the results described in these two tables, it can be seen that the upper participants categorized polysemous senses in more systematic ways, similar to the dictionary definitions, but the lower participants' mental lexicons were not as developed as those of the upper participants. This resulted in the categorizations being less systematic and far from the dictionary classifications.

Let us compare the results of this study and **Figure 1**. What we have found is that the EFL students with a large vocabulary size made categories more similar to those in the dictionary. Thus, as vocabulary size increases, the structure of mental lexicon develops and becomes more similar to a linguistic specialists' perception. That is, the upper group created categories closer to A, B, C, or D in

Table 5. Morey & Agresti's (1984) adjusted rand index comparing participants' classification and dictionary's classification.

Target Word	<i>n</i> of senses	Group	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>	Effect size (<i>d</i>) [95%CI]
go	3	Upper	27	0.80	0.17	2.16	51	0.035	0.59 [0.03, 1.16]
		Lower	26	0.68	0.22				
look	3	Upper	27	0.79	0.16	3.79	51	0.000	1.04 [0.45, 1.63]
		Lower	26	0.59	0.22				
time	3	Upper	27	0.56	0.13	1.63	51	0.109	0.45 [-0.11, 1.01]
		Lower	26	0.50	0.16				
take	4	Upper	27	0.56	0.14	3.41	49	0.001	0.96 [0.36, 1.55]
		Lower	24	0.43	0.12				
get	5	Upper	27	0.52	0.11	2.33	50	0.024	0.65 [0.08, 1.22]
		Lower	25	0.45	0.10				
have	7	Upper	27	0.49	0.12	1.09	50	0.283	0.30 [-0.26, 0.85]
		Lower	25	0.45	0.13				

If we adopt Bonferroni's correction, the critical *p*-value becomes 0.008 (0.05 divided by 3).

Table 6. Morey & Agresti's (1984) adjusted rand index comparing classifications of each participant and intra-group.

Target Word	<i>n</i> of senses	Group	<i>n</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>	Effect size (<i>d</i>) [95%CI]
go	3	Upper	27	0.70	0.11	4.11	51	0.000	1.13 [0.53, 1.72]
		Lower	26	0.58	0.10				
look	3	Upper	27	0.69	0.10	7.85	51	0.000	2.16 [1.46, 2.85]
		Lower	26	0.48	0.09				
time	3	Upper	27	0.52	0.07	1.93	51	0.059	0.53 [-0.03, 1.09]
		Lower	26	0.47	0.10				
take	4	Upper	27	0.58	0.12	2.17	49	0.035	0.61 [0.03, 1.19]
		Lower	24	0.51	0.11				
get	5	Upper	27	0.53	0.08	7.54	50	0.000	2.09 [1.40, 2.78]
		Lower	25	0.38	0.06				
have	7	Upper	27	0.45	0.08	3.59	50	0.001	1.00 [0.40, 1.59]
		Lower	25	0.37	0.08				

If we adopt Bonferroni's correction, the critical *p*-value becomes 0.008 (0.05 divided by 3).

Figure 1 than the lower group. More specifically, the upper group tended to put A_1 , A_2 , and A_3 , which were the same in each target context in this study, into the same category. However, considering that native speakers still produced a different number of senses from the dictionary's definition (Lin & Ahrens, 2010), categorization in the mental lexicon might not be able to be perfectly the same as

in the dictionary. Still, it becomes closer to the dictionary's definition as the proficiency develops.

Furthermore, the upper group's categories were more converged than the lower groups' classification, meaning that the participants in the upper group classified the target contexts similarly; on the other hand, the lower group's classifications had more variation, even though both groups answered regarding the target context and showing understanding to the same degree (see **Table 3**). Two possibilities would be considered based on *Aitchison's (2003)* model wherein the development of knowledge of word meaning started from labeling, and then went to packaging (categorization) and finally network building.

The first possibility is that the difference between the two groups was caused in the packaging process; that is, although both groups understood the context well and had already reached the labeling stage, the lower group's packaging knowledge was not as developed as the upper group's. The results of this study thus make up one piece of evidence to support the idea that knowing the sense of a polysemous word requires different knowledge or ability than does making a categorization. Another possibility is that the difference between the two groups was caused in the labeling process, and that might affect the following packaging process. In the current study, since participants self-assessed how well they understood each sentence, there is a possibility that the participants in the lower group thought they understood each context well, when in actuality their comprehension might be more ambiguous than that of the participants in the upper group. Both possibilities are plausible in an L2 context, but one thing to note is that even the lower-level students in this study knew 4635 words on average based on the VST. Thus, it is true that their knowledge about very basic polysemy was still underway, but the former possibility would be more reasonable. These facts indicate that teachers should focus not only on connecting L1 senses and L2 target polysemous words (i.e., labeling), but also on how each sense of polysemy can be categorized by showing example sentences including the target word (i.e., packaging). By doing so, students can be more likely to organize their mental lexicon, which will come to closer to the dictionary's categorization.

4. Conclusion

This study investigated how Japanese EFL learners classified the senses of basic polysemy. The answer to RQ1 ("What kinds of differences were found between the sense categorizations made by the dictionary and those made by Japanese EFL learners?") is that the participants tended to make more sense groups than the dictionary's categorization. That is, learners make more minute categorizations, which are beyond the linguistic experts' classifications described in dictionaries. Next, the answer to RQ2 ("Are there any differences in sense categorization in relation to the English vocabulary level of Japanese EFL learners?") is that the number of categories each group made did not differ, but the structure

of the sense groups was different. To be more specific, the upper students with larger vocabulary sizes categorized senses more similarly to the dictionary's classification than the lower students with smaller vocabulary sizes, and the upper-level students' categorization was more converged compared to the lower group. As the vocabulary knowledge increased, the mental lexicon became more structured and students could define and group the words according to the same senses described in the dictionary. Even though neither group of students had difficulty in understanding the target contexts, differences were found in the categorization of the senses of polysemy between the two groups. This fact indicates that acquiring meaning requires different skills than for categorization or networking. However, as Lin and Ahrens (2010) showed, it is possible that categorization in the mental lexicon never completely matches the dictionary's classification, no matter how proficient the learners become. Human beings might classify the senses of polysemy not only from a theoretical point of view, but also based on their personal experience, the input about the polysemy they have received, or their instinct (no definable reason at all). It is necessary to investigate the reasons for making sense categorizations, which is beyond the scope of this study.

Since few studies using polysemy have been conducted in an L2 context, the research on the structure of L2 lexical networks is still in development. Furthermore, this study only focused on part of this structure; that is, how the nodes of six basic polysemous words were constructed. In order to fully examine the structure of the mental lexicon, it would be necessary to include more words as well as more senses, although that would impose more of a burden on the participants. However, the current study is insightful in showing that even though L2 learners seem to acquire some basic words and understand their senses (the labeling stage), a difference appears in the ability to categorize the senses (the packaging stage). Furthermore, this knowledge can be gradually developed as L2 learners' proficiency develops. Further research is necessary to replicate the results in this study and generalize the process of understanding polysemous words in L2 learners' mental lexicon, especially considering the sequence of labeling, packaging and networking stages. Such research could finally clarify how a lexical network is constructed and developed in a learner's mental lexicon as well as how teachers can encourage learners to develop this network through educational mediation.

Acknowledgements

We thank Dr. Nobuhiko Akamatsu for giving valuable comments as well as helping with the analysis.

References

- Aitchison, J. (2003). *Words in the Mind: An Introduction to the Mental Lexicon*. UK: Blackwell.

- Azuma, T., & van Orden, G. C. (1997). Why SAFE Is Better than FAST: The Relatedness of a Word's Meanings Affects Lexical Decision Times. *Journal of Memory and Language*, *36*, 484-504. <https://doi.org/10.1006/jmla.1997.2502>
- Brown, S. W. (2008). Polysemy in the Mental Lexicon. *Colorado Research in Linguistics*, *21*, 1-12. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.518.1533&rep=rep1&type=pdf>
- Collins, A. M., & Loftus, E. G. (1975). A Spreading-Activation Theory of Semantic Processing. *Psychological Review*, *82*, 407-428. <https://doi.org/10.1037/0033-295X.82.6.407>
- Collins, A. M., & Quillian, M. R. (1969). Retrieval Time from Semantic Memory. *Journal of Verbal Learning and Verbal Behavior*, *82*, 240-247. [https://doi.org/10.1016/S0022-5371\(69\)80069-1](https://doi.org/10.1016/S0022-5371(69)80069-1)
- Foraker, S., & Murphy, G. L. (2012). Polysemy in Sentence Comprehension: Effects of Meaning Dominance. *Journal of Memory and Language*, *67*, 407-425. <https://doi.org/10.1016/j.jml.2012.07.010>
- Frazier, L., & Rayner, K. (1990). Taking on Semantic Commitments: Processing Multiple Meanings vs. Multiple Senses. *Journal of Memory and Language*, *29*, 181-200. [https://doi.org/10.1016/0749-596X\(90\)90071-7](https://doi.org/10.1016/0749-596X(90)90071-7)
- Hoshino, Y. (2016). Which Meanings of Basic Words Appear in English Reading Tests with Various Difficulties?—Focusing on Polysemy. *Annual Review of English Language Education in Japan*, *27*, 33-48. https://doi.org/10.20581/arele.27.0_33
- Jiang, N. (2000). Lexical Representation and Development in a Second Language. *Applied Linguistics*, *21*, 47-77. <https://doi.org/10.1093/applin/21.1.47>
- Klepousniotou, E. (2002). The Processing of Lexical Ambiguity: Homonymy and Polysemy in the Mental Lexicon. *Brain and Language*, *81*, 205-223. <https://doi.org/10.1006/brln.2001.2518>
- Klepousniotou, E., Titone, D., & Romero, C. (2008). Making Sense of Word Senses: The Comprehension of Polysemy Depends on Sense Overlap. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 1534-1543. <https://doi.org/10.1037/a0013012>
- Langacker, R. W. (1999). *Grammar and Conceptualization*. New York: Mouton de Gruyter. <https://doi.org/10.1515/9783110800524>
- Lin, C., & Ahrens, K. (2010). Ambiguity Advantage Revisited: Two Meanings Are Better than One When Accessing Chinese Nouns. *Journal of Psycholinguistic Research*, *39*, 1-19. <https://doi.org/10.1007/s10936-009-9120-8>
- Mizumoto, A., & Takeuchi, O. (2008). Basics and Considerations for Reporting Effect Sizes in Research Papers. *Studies in English Language Teaching*, *31*, 57-66. http://www.mizumot.com/files/EffectSize_KELES31.pdf
- Morey, L. C., & Agresti, A. (1984). The Measurement of Classification Agreement: An Adjustment to the Rand Statistics for Chance Agreement. *Educational and Psychological Measurement*, *44*, 33-37. <https://doi.org/10.1177/0013164484441003>
- Nissen, H. B., & Henriksen, B. (2006). Word Class Influence on Word Association Test Results. *International Journal of Applied Linguistics*, *16*, 389-408. <https://doi.org/10.1111/j.1473-4192.2006.00124.x>
- Rodd, J., Gaskell, G., & Marslen-Wilson, W. (2002). Making Sense of Semantic Ambiguity: Semantic Competition in Lexical Access. *Journal of Memory and Language*, *46*, 245-266. <https://doi.org/10.1006/jmla.2001.2810>
- Seto, K. (2007). *Shogakukan dictionary of English lexical polysemy*. Tokyo: Shogakukan.

- Tanaka, S. (2016). *Eigo wo tsukaikonasu tameno jissen-teki gakushu-hou my English no susume [Practical Learning Method for Mastering English]*. Tokyo: Taishukan.
- Wolter, B. (2001). Comparing the L1 and L2 Mental Lexicon. *Studies in Second Language Acquisition*, 23, 41-69. <https://doi.org/10.1017/S0272263101001024>
- Wolter, B. (2006). Lexical Network Structures and L2 Vocabulary Acquisition: The Role of L1 Lexical/Conceptual Knowledge. *Applied Linguistics*, 27, 741-747. <https://doi.org/10.1093/applin/aml036>
- Zareva, A. (2007). Structure of the Second Language Mental Lexicon: How Does It Compare to Native Speakers' Lexical Organization? *Second Language Research*, 23, 123-153. <https://doi.org/10.1177/0267658307076543>