

A Study on Lexical Sense Relations from the Perspective of Vocabulary Breadth and Word Frequency

Xiaoli Cheng

Chongqing Normal University, Chongqing, China

Abstract—In second language (L2) vocabulary acquisition, breadth and depth of vocabulary knowledge are two indispensable components that interrelate with each other to a substantial extent. Breadth of vocabulary is actually vocabulary size. Lexical sense relations, part and parcel of depth of vocabulary, are reported to be able to facilitate the mastery of L2 words. Word frequency concerns the familiarity of words. The present research intends to make a synthetical analysis of the development of L2 word meaning under the influence of vocabulary size and word frequency in classroom settings. The merging of qualitative and quantitative aspects of words is to describe exhaustively how students fit the words into their mind, and to provide some pedagogical implications to L2 vocabulary teaching and learning.

Index Terms—depth, breadth, word frequency, lexical sense relations

I. INTRODUCTION

Depth of L2 vocabulary knowledge never fails to attract attention in the field of L2 vocabulary research. According to Qian (1990, p.283), depth of vocabulary knowledge is about a learner's level of knowledge of different aspects of a word, or how well the learner knows a word. Linguists in the literature once provided various understandings to the content of it (Laufer, 1997; Nation, 2001). Soderman (1989) carried out an experiment to testify that learners with different language proficiency would demonstrate diversified word relation associations. In China, Gui (1992) made an experiment to test the relationships between Chinese and English mental lexicon. Liu (2001) conducted empirical research to delineate the development of L2 word meaning and word affixes. The present research, in light of the aforementioned achievements, attempts to provide a much in-depth account of the development of L2 word meaning relations under the influence of vocabulary size and word frequency. The purpose is to arouse Chinese L2 learners' awareness to word meaning relations in different learning stages and to provide some pedagogical implications to L2 vocabulary acquisition.

II. LITERATURE REVIEW

A. *Word Meaning, Vocabulary Size and Word Frequency*

Depth of vocabulary knowledge can be portrayed as a hierarchical structure consisting of layers of contents that are of different importance and extending from superficial to deep level processing in our mind. L2 word meaning, a fundamental component of this hierarchy, is thought to encompass two basic parts: reference and sense. The former concerns how language hooks onto the world, that is, how words are used to represent the entities or the occurrences in the outside world, while the latter specifies the semantic links between elements within the vocabulary system (Saeed, 2001). The present research is just devoted to the exploration of sense relations between words.

Vocabulary size and word frequency are two quantitative features of words, each indicating the level of L2 learners and the familiarity of L2 words respectively. Vocabulary size, which is also formally termed as breadth of vocabulary, refers to "the number of words for which a learner has at least some minimum knowledge of meaning" (Qian, 1999, p.283). Therefore, it is also treated as a mirror to show how educated, intelligent or well-read a person is. Word frequency denotes how often the word occurs in normal use of the language (Schmitt & McCarthy, 1997). According to Nation (1990), there are high-frequency words, university words, technical words and low-frequency words. With the storage of about 3000 high-frequency and university words, an L2 learner is supposed to be able to cover 90% to 95% of any text and can safely transfer his first language (L1) reading strategies into L2 learning.

Breadth and depth of L2 vocabulary are not two isolated parts. They are the quantitative and qualitative facets of vocabulary knowledge. Empirical research testifies that there exist strong correlations between breadth and depth of vocabulary. The inadequacy of one may impair the development of another (Qian, 1999). Due to this close interdependence between them, it must be feasible and beneficial for us to explore the development of L2 lexical sense relations under the influence of vocabulary size. As to word frequency, being an extrinsic characteristic of words, how the change of it would influence the intrinsic representation of words may also pose an interesting and challenging task for language researchers.

B. *Sense Relations of Words*

Saeed (2001) observed that it may be more accurate to think of the lexicon as a network. The organizational principle of this network is not specified so far, but one point is quite clear that this is not a network without order. In fact, it is one full of interconnected elements or nodes, which are connected to one another by virtue of having various relations (Carrol, 2000, p.110). Paradigmatic and syntagmatic relations of words are commonly regarded to be two dominant lexical sense relations. The former is also called substitutional relation, which holds between intersubstitutable members of the same grammatical category (Lyons, 1995, p.124). To place words that contain paradigmatic relations in the context of sentences, they can substitute each other without changing the syntactical structure or violating grammatical rules. Syntagmatic relation is also termed combinational relation, which exists typically, though not necessarily, between expressions of different grammatical categories that can be put together in grammatically well-formed combinations or constructions (*ibid.*), say, between nouns and adjectives, verbs and nouns, or verbs and adverbs etc.

However, relations between words are more than sense relations. Clang responses (responses related to stimuli in phonological terms only) are also included (Marechal, 1995). Therefore, in order to present an exhaustive account of the development of L2 lexical sense relations, I introduce formal relations into my experiment, which refer to those similarities between words in terms of word form, such as similarities in spelling, pronunciation or derivation. Therefore, in this article, I totally adopt three kinds of lexical relations: paradigmatic relation, syntagmatic relation and formal relation. The first two are sense relations between words, while the third one includes spelling & pronunciation relation and derivational relation.

III. RESEARCH DESIGN

A. *Research Questions*

According to level-of-processing theory, meaning should be ascribed to deep-level-processing, which will probably bring about long-time or even permanent memory (Craik&Lockart, 1972). To Chinese L2 learners, after years of English learning and with the assistance from their L1 meaning system, will they approach L2 words from the perspective of lexical sense relations? Will this meaning-oriented processing change with the alteration of L2 learners' breadth of L2 vocabulary? Will the frequency or familiarity of L2 words affect L2 learners' processing mode?

B. *Research Methods*

In the experiment, Word Association Test, which is the oldest method psychologists have for studying semantic relations, is employed. This test is invented by Sir Francis Galton and has been widely used by many language researchers. The exact procedure of this test is as follows: first, researchers provide some stimuli words to their subjects, who are required to write down the words which strike their mind as the first response. Then the relations between the stimuli words and the response words are analyzed and the organizational principle of L2 learners' mental lexicon can be tentatively analyzed and uncovered.

In terms of stimuli words, considering the vocabulary size of my research subjects, the first 5000 frequently used words are borrowed from an influential Chinese database entitled *An English Corpus Based on Chinese Learner* (Gui & Yang, 2003). Then these words are divided into five groups according to their word frequency. From each group, one word every 100 words is chosen. Finally 50 words are obtained and treated as my research targets.

In the controlling of test environment, in order to avoid slapdash associations and associations after careful thinking, the word association test is implemented in class and time restriction is set on the subjects. Besides, no explanations or hints about the stimuli words are offered to the subjects. After the experiment, in order to guarantee the consistency and reliability of data analysis, ten students' association papers are randomly chosen and independently analyzed by researchers. The analysis of each researcher is compared, and differences are discussed in order to sort out a norm for future use. The analysis of all the data lasts for about one week.

C. *Research Subjects*

The 140 subjects in my experiment are the first and the third year college students who major in English in a Chinese university. Students in grade one have been in college for about three months. According to the requirements of the teaching curriculum (2007) for English majors in China, their average breadth of L2 receptive vocabulary reaches about 2000-3000 words, with about 1000-1200 productive words. The average receptive vocabulary size of the third-year students is about 5000-6000 words, including about 2000-2500 productive words.

IV. DATA ANALYSIS

The experiment centers on four association types: paradigmatic, syntagmatic, derivational, spelling and pronunciation associations. There may exist other types of individualized, slapdash associations. Because they denote nothing about meaning or formal characteristics of words, they are not included in this thesis. As is indicated, paradigmatic and syntagmatic associations are deep-level meaning associations. In contrast, the other two associations belong to surface-level processing. Table 1 provides the number of associations in each level.

TABLE 1.
THE NUMBERS OF DIFFERENT ASSOCIATIONS IN EACH LEVEL

	grade	paradigmatic relation	syntagmatic relation	total	spelling & pronunciation relation	derivational relation	total
1 st 1000	One	234	62	296	238	142	380
	Three	434	168	602	28	48	76
2 nd 1000	One	274	86	360	200	136	336
	Three	412	212	624	18	38	56
3 rd 1000	One	150	50	200	350	122	472
	Three	300	238	538	52	78	130
4 th 1000	One	204	76	280	158	242	400
	Three	388	188	576	24	78	130
5 th 1000	One	210	56	266	264	148	412
	three	300	182	482	70	122	192

A. *Word Meaning and Vocabulary Size*

From Table 1, in each level, the summation of meaning associations of the subjects is generally larger than that of their formal associations, which leaves us an impression that to process L2 words in terms of meaning seems to be a much favored choice among Chinese intermediate and advanced L2 learners. This observation seems to render a much satisfactory answer to the first research question. However, if we continue to analyze the data with enough concentration and patience, one fact shouldn't be ignored, that is, in each level, most of the meaning associations are demonstrated by the third-year students, although formal associations do exist in their responses. In contrast, the responses of the first-year students are dominated mainly by formal associations, even to words in the first level. How can we account for this phenomenon? Can we safely conclude that vocabulary size does affect response types? And how?

Words are not isolated from each other. With the accumulation of L2 words, the relations between them become more compact and also more diverse. Gradually, a semantic network of L2 words comes into existence and the formal connections between words begin to give way to meaning associations. In the experiment, the performance of the third-year students is encouraging. In each level, meaning associations surpass formal associations enormously, even to words in the fifth level. As a sharp contrast, the response types of those freshmen are not so satisfying. Although there still exist quite a number of meaning associations in each level, they lag far behind the formal associations, even to words in the first level. This contrast between meaning and formal associations to students in different grades is, in my opinion, due to their divergence in vocabulary size. All the chosen words are within the vocabulary size of the students in grade three, while many of them are beyond the competence of the first-year students. With a limited semantic network, those freshmen have no choice but to turn to formal connections between words. In a word, breadth of L2 vocabulary does affect L2 learners' way of processing L2 words. We can hopefully predict that with the expansion of vocabulary storage, L2 learners will gradually become inclined to deal with L2 words by way of lexical sense relations. The more words they memorize, the more meaning associations they will demonstrate.

Also from Table 1, to the sense relations of words, both groups of students display the majority of paradigmatic associations, that is, students tend to associate synonyms, antonyms or words of the same part of speech to the stimuli words. The proportion of syntagmatic associations to the total meaning associations of the first-year students is smaller than that of the third-year students. To the students in grade three, their syntagmatic associations don't diminish but to continue to increase. I think this phenomenon can be reasonably explained though it is inconsistent with Soderman's (1989) claim that like native speakers, less proficient L2 learners will also demonstrate more syntagmatic associations than adept L2 learners. As far as I am concerned, syntagmatic association is a manifestation of students' ability to combine or collocate words grammatically. It is a preparatory capacity for the emergence of well-formed sentences or larger discourse. Therefore, there is no reason to predict that syntagmatic association will decrease with the improvement of vocabulary size. Instead, it may also exhibit an increase, though a mild one. If so, what matters most in future research is not about the numbers of syntagmatic association, but the types of it, that is, what are the words which are usually put together by developed or less developed students?

B. *Lexical Sense Relations and Word Frequency*

With regard to the relations between lexical sense relations and word frequency, due to the poor vocabulary size of the first-year students and their irregular performances in the test, I employ only the performance of the third-year students as my target of research. Table 2 presents the average (the total number of each lexical association in different levels versus the number of subjects who hold that specific association) and the standard variation of each response type in the five levels.

TABLE 2.
THE AVERAGE AND THE STANDARD VARIATION OF THE DATA

	Paradigmatic relation		Syntagmatic relation		Spelling and Pronunciation		Derivation	
	A.	S.	A.	S.	A.	S.	A.	S.
1st 1000	6.20	1.83	2.90	1.23	1.40	.96	1.41	.62
2nd 1000	6.05	1.61	3.53	.47	1.12	.36	1.58	.52
3rd 1000	4.29	1.78	3.50	1.85	1.37	.69	1.77	.32
4th 1000	5.39	2.11	3.03	1.60	1.50	1.07	2.29	1.90
5th 1000	4.54	1.68	2.84	2.05	1.59	.91	2.44	1.39

Notes: A=Average S=Standard variation

On the basis of the means in table 2, a graph is drawn to illustrate the changing route of each response type (Figure 1). From Figure 1, paradigmatic association is still the most frequent response type. Paying close attention to its changing route in each level, we can detect that the first 2000 words possess the most paradigmatic associations. Although responses to the third, the fourth and the fifth level of words fluctuate, the changing route of paradigmatic associations to word frequency is, in general, declining. With the increase of the unfamiliarity of L2 words, paradigmatic associations decrease accordingly. When the stimuli words are beyond the vocabulary size of L2 learners, paradigmatic associations will finally give way to formal associations again.

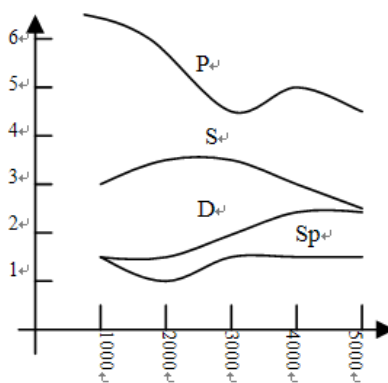


Figure 1: The changing route of each response type
Note: P=paradigmatic S=syntagmatic D=derivational Sp=spelling & pronuncional

In terms of syntagmatic association, its alteration is mild. According to Figure 1, words in the first and fifth level demonstrate the least syntagmatic associations, while the second, the third and the fourth 1000 words contain more syntagmatic associations, with those of the third 1000 words coming up to a peak. The curve of syntagmatic associations is puzzling, but if we take the change of the other associations into consideration, the whole picture becomes clearer. The associations to high and low frequency words are dominated by paradigmatic and formal associations respectively. This phenomenon confirms a fact that Chinese L2 learners are not sensitive to syntagmatic connections between words even when those words are within their productive vocabulary size. Therefore, in practical use of L2, they are found to be incompetent in the mastery of the modifications and collocations of L2 words.

The changing route of derivational associations comprises an increasing slope. The amount of derivational associations to the first 1000 words lags far behind paradigmatic and syntagmatic associations. But to the fifth 1000 words, the number of derivational association begins to approximate that of syntagmatic association. According to its developing tendency, it may finally get the upper hand over syntagmatic even paradigmatic association in the end when L2 words are far beyond the vocabulary size of L2 learners. L2 learners' inclination to derivational associations is possibly due to the emphasis placed on them in L2 teaching and learning.

Spelling and pronuncional associations, to the third-year students in my experiment, contain the least number. Its changing route is mild too, although it still improves with the increase of word frequency. To adult L2 learners, within their vocabulary size, with the positive transfer of the L1 meaning system, spelling and pronuncional associations lose their dominance completely.

The standard variation of different response types in each level indicates that to words with high frequency, the divergence of associations among students is small. However, when word frequency is beyond the fourth 1000 words, this divergence becomes conspicuous, which may disclose the probability that students' response types will become more and more diverse with the increase of word frequency.

V. CONCLUSIONS

On the basis of the experiment, we can tentatively draw the conclusion that Chinese adult L2 learners have the tendency to approach L2 words semantically, and vocabulary size and word frequency do affect L2 learners' association types. When words are within their command, students are generally inclined to connect words in terms of sense

relations, especially those with paradigmatic relations, such as synonymy, antonymy, hyponymy or part-whole relations. Although syntagmatic associations are not salient, they also develop mildly. Formal associations are usually resorted to when the stimuli words are out of the students' vocabulary size. As to word frequency, the first 5000 words in general trigger more sense relations than formal relation of words. With the dropping of word familiarity, formal associations may overtake the meaning relations. Besides, there seems to exist a threshold to breadth of vocabulary. Below it, students are likely to demonstrate formal associations, even to very high frequency words. Put it in another way, students should grasp a minimum number of words so as to build a necessary semantic network, which may help to store words in the long-term memory. Therefore, in L2 vocabulary acquisition, learners should build a solid, sufficient word storage to guarantee an expansive semantic network, otherwise meaning connections between words may become quite random.

At present, many Chinese learners of English pay keen interest in enlarging vocabulary size. However, the acquisition of the depth of vocabulary knowledge usually receives less attention, let alone combining these two efforts together. As is testified, L2 learners become sensitive to word meaning relations only when they have accumulated sufficient words. Therefore, in the preliminary stage of L2 vocabulary teaching and learning, introducing too many words, which are related with the target word semantically, may bring about unexpected burden and confusion. A wiser practice is still to embed target words into contexts to strengthen understanding of their meaning and use.

Placing words into contexts is also facilitative for students to observe and analyze the co-text of word using, that is to cultivate students' competence to produce well-formed collocations, phrases and sentences. From the experiment, Chinese L2 English learners are less adept in syntagmatic associations. Therefore, in classroom teaching and learning, combinational relations among words from different categories can be accentuated. If possibly, some prefabricated constructions can also be stressed and memorized. To unfamiliar words and before reaching the threshold of vocabulary size, although students are subject to formal associations, it doesn't follow that association in terms of pronunciation, spelling or derivation should be encouraged in classroom vocabulary teaching and learning, for formal associations may also produce, if not more, the same amount of confusing learning burdens.

As to L2 word frequency, it receives increasing stress and attention in recent years. However, the implementation of its achievements in classroom settings has not been extensively realized. At the end of each textbook, vocabulary is usually arranged alphabetically. In each unit, new words are conventionally presented on the basis of their order of appearance in the passage. Many middle school English teachers haven't got any sense of word frequency in their foreign language teaching. Fortunately, linguists and educators in China are making unprecedented contributions to this barren land. Empirical research is conducted and data bases are being built. Gui Shichun and Yang Huizhong (2003) are one of them. Their book entitled *An English Corpus Based on Chinese Learners* provides a detailed and exhaustive description of the frequency of English words on the basis of Chinese L2 learners. This endeavor may arouse more and more awareness to L2 vocabulary research among foreign language teachers, textbook compilers, etc.

The present study carried out an empirical research in order to describe how the relations between words develop with the improvement of language ability and the change of word frequency. The result is enlightening although there exist much room for future refinement, for example, research subjects are from the same university, and the number of students is limited, which may affect the generalizability of the present research. Besides, the experiment doesn't differentiate the effects of receptive and productive vocabulary size on depth of word knowledge, although it is mentioned in the paper. Last but not least, the lexical relations between words are analyzed within the abstract vocabulary system, ignoring those social factors such as the characteristics of learners or the learning environment of words. With the development of such disciplines as cognitive semantics or construction grammar, the combination of social factors into research may promise a brighter future.

REFERENCES

- [1] Carroll, David W. (2000). *Psychology of Language*. Beijing: Foreign Language Teaching and Research Press.
- [2] Craik & Lockhart. (1972). Level of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11, 671-684.
- [3] Gaodeng xuexiao yingyu zhuan ye jichu jieduan yingyu jiaoxue dagang [English Teaching Syllabus for English Majors in the Elementary Stage of Higher Education]. (2007). Shanghai: Shanghai Foreign Language Educational Press.
- [4] Gui, Shichun. (1992). Yuyan yu waiyu xuexi [Language and language learning]. *Foreign Language Teaching and Research* 4, 2-9, 80.
- [5] Gui, Shichun & Huizhong Yang. (2003). *Zhongguo xuexizhe yingyu yuliao ku* [An English Corpus Based on Chinese Learners]. Shanghai: Shanghai Foreign Language Educational Press.
- [6] Laufer, B. (1997). The lexical plight in second language reading. In *Second Language Vocabulary Acquisition*. Eds. Coady & Huckin. Cambridge: Cambridge University Press.
- [7] Liu, Shaolong. (2001). Eryu cihui shendu xide ji fazhan tezheng [Depth of L2 vocabulary acquisition and its developmental features]. *Foreign Language Teaching and Research* 6, 436-441, 480.
- [8] Lyons, J. (1995). *Linguistic Semantics: An Introduction*. Cambridge: Cambridge University Press.
- [9] Marechal, C. (1995). *The Bilingual Lexicon: Study of French and English Word Association Responses of Advanced Learners of French*. M. Phil. Dissertation. University of Dublin.
- [10] Nation, I.S.P. (1990). *Teaching and Learning Vocabulary*. New York: Newbury House Publishers.
- [11] Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.

- [12] Qian, David D. (1999). Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review* 56, 282-307.
- [13] Saeed, J.I. (2001). *Semantics*. Beijing: Foreign Language Teaching and Research Press.
- [14] Schmitt, Norbert & Michael McCarthy. (1997). *Vocabulary: Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.
- [15] Soderman, T. (1989). Word associations of Foreign language learners and native speakers—A shift in response type and its relevance for a theory of lexical development. *Scandinavian Working Papers on Bilingualism* 8, 114-121.



Xiaoli Cheng, received her MA degree from Sichuan International Studies University in Chongqing, China. She also got a diploma of teaching methodology from National Institute of Education of Nanyang Technological University in Singapore. Now she is a lecturer at the School of Foreign Language and Literature, Chongqing Normal University. Her research interests include second language acquisition, teaching methodology, pragmatics, etc.