

What Insights does Research on Vocabulary in Reading Provide for the Teaching of Reading?

Kazuo Nakajima

1. Introduction

It has long been acknowledged that vocabulary knowledge is the most important factor affecting reading comprehension. For example, Chall (1958: 157; cited in Nation and Coady, 1988: 97) claims: "once a vocabulary measure is included in a prediction formula, sentence structure does not add very much to the prediction." Accordingly, there have been a number of researches aimed at examining the relationship between the vocabulary knowledge and comprehension in both L1 and L2. On the other hand, it has also been widely accepted that reading facilitates vocabulary learning and there have been considerable numbers of researches supporting this belief (although some of them suggest some limitations on the role reading plays in vocabulary learning).

Although the two issues are closely related to each other, and some of the researches address the both at the same time, I would like to review them in the two different contexts; 1) How the vocabulary knowledge (and strategies) contributes to reading proficiency, and 2) How reading activities facilitate vocabulary learning. In the course of review, I would like to consider implications these researches have for teaching reading in EFL classrooms.

2. Vocabulary Knowledge and Reading Proficiency

2-1. Vocabulary Size and Lexical Coverage in a Text

The question of how much vocabulary a learner needs to have in order to read in L2 satisfactorily has been addressed by many researchers. For example, Laufer (1992) examined the correlation between the reading comprehension scores and the vocabulary size levels using university students as subjects, and concluded that the level of 3000 word families was the turning point (She drew that conclusion from the fact that there were more students who reached the minimal comprehension score required by the course than those who failed above that level, and there was a statistically significant gap between 2000 level and 3000 level.). Schmitt (2000: 143), citing Nation and Waring (1997), also claims that 3000—5000 word families would be consensus for the first access to authentic texts.

In terms of optimal lexical coverage (how many percentage of the words in a text must be known for successful comprehension), or the density of unknown words, there seems to be considerable difference according to researchers presumably because of other factors involved (e.g. length of the passage, as suggested in Carter and McCarthy, 1988: 99). Nation and Coady (op. cit.), after reviewing several researches concerning the lexical density, suggest to follow West's (1941) guideline of 2 percent of unknown words in a text. Laufer (1992: 126-127) cites Deville (1985) and Laufer's (1989) claim that 95 percent of lexical coverage is required for academic level reading.

It should be noted that vocabulary size and lexical

coverage are closely related to each other in that the larger vocabulary size a learner has, the fewer unknown words s/he would find in a text. Therefore, it might be said that it is lexical coverage rather than vocabulary size that directly affects the readability of a certain text a learner tries to comprehend. It implies usefulness of graded readers because they enable learners with relatively small size of vocabulary to actually "read" by controlling words used in the text according to their frequency levels. (Many researchers also emphasize the usefulness of graded readers for vocabulary 'learning', but we will consider their claims later.)

If there is a certain critical level of vocabulary size for reading authentic materials as some of the researches suggest, one of the major goals of reading instruction in EFL classrooms would be to have the students acquire that level of vocabulary.

2-2. Depth of Vocabulary Knowledge

The size of vocabulary is sometimes termed 'breadth' of vocabulary knowledge (e.g. Anderson and Freebody, 1981; cited in Read, 1993: 357). It is the matter of how many words a learner knows. On the other hand, it is also important to know how well a learner knows a particular word; that is, 'depth' of vocabulary knowledge. It seems that the 'depth' of vocabulary knowledge has been paid less attention to compared to the 'breadth' in the field of L2 research. As Nation (1990: 30-31) explains, knowing a word entails several aspects. He classifies them into four categories; form (spoken / written), position (grammatical patterns / collocations), function

(frequency / appropriateness) and meaning (concept / associations). Schmitt (2000: 5), basically adopting Nation's framework, asserts that acquisition of each aspect is gradual one: "Each of word-knowledge types is likely to be learned in gradual manner, but some may develop later than others and at different rates." If that is the case, it follows that it is likely that a learner's knowledge of a particular word is somewhere between absolute zero and complete mastery at a certain point of time (We will discuss this 'incremental' nature of vocabulary learning later in this paper.).

In recent years, there have been several attempts to devise measures to assess the depths of vocabulary knowledge of learners. One of them is Read's (1993) 'depth of vocabulary knowledge test' utilizing word association. He (op. cit: 358) justifies the choice of word association as a measure of the depth of vocabulary knowledge citing findings of previous researches (Meara, 1980 etc.) that patterns of word association produced by L2 learners tend to be unstable compared to remarkably stable patterns of native speakers.

Empirical researches investigating the relationship between the depth of vocabulary knowledge and reading comprehension have been sparse. However, Qian's (1999) study revealed that the depth of vocabulary knowledge, as well as the breadth of vocabulary knowledge, plays an important role in predicting reading comprehension. Noro (2002), basically replicating Qian's research, made additional findings that importance of the depth of vocabulary knowledge increases as the vocabulary size of learners approaches 3000 word families.

The result of these researches implies that it is necessary not only to expand the vocabulary size of learners but to improve the quality of their vocabulary knowledge especially for higher level learners.

2-3. Lexical Inference in Reading

It is unlikely especially for L2 learners to find no unfamiliar words on a page while reading authentic materials. Therefore, strategies for guessing unknown words from context have been considered crucial reading skills to be a good reader.

Nation and Coady (1988: 104-105) suggests the following five-step guessing strategy:

1. Finding part of speech of the unknown word.
2. Looking at the immediate context of the unknown word and simplifying this context if necessary.
3. Looking at the wider context of the unknown word. This means looking at the relationship between the clause containing the unknown word and surrounding clauses and sentences.
4. Guessing the meaning of the unknown word.
5. Checking that the guess was correct.

His suggestion is basically looking at the word itself and local context first, then, looking for clues in wider (global) context. However, he (op. cit: 107) suggests to reserve the use of word forms until the checking step based on the findings of Looby (1939), Gibbons (1949), Haynes (1984) and Bensoussan and Laufer (1984) that the use of word forms for guessing in the first place often leads to wrong guess. Coady also notes that

learners often misinterpret the context so that it fits to the wrong meaning of the words inferred from word forms. The same problem is reported by Huckin and Bloch (1993) ('mistaken ID') and Laufer (1997) ('deceptive transparency').

There could be variety of ways to teach learners efficient strategies for guessing the meaning of unknown words. In her case study, Hosenfeld (1984) tried two different ways of strategy instruction to her subjects. She conducted inductive awareness-raising activities (where the subject was asked to compare her own strategies with those of a proficient reader) and deductive instruction through dialogue, and succeeded in improving their use of strategies in both the cases.

2-3-1. Limitations on Lexical Inference

Laufer (1997:27-30), while admitting some value of guessing strategies, casts doubt on its usefulness, noting a variety of factors that interfere with the guessing attempts of the reader, such as 'nonexistent contextual clues', 'unusable contextual clues', 'misleading and partial contextual clues' and 'suppressed clues'. It is certain that learners do benefit from word-guessing strategy instruction, but as Laufer suggests, it is not applicable to every situation. It might be wise for us to teach learners efficient ways of using outer resources such as dictionaries as well as inferencing strategies.

There are several other researches based on 'threshold hypothesis' which implies limitations on the efficacy of lexical inference strategies. Alderson (1984), after reviewing a number of relevant empirical researches, concludes that a

certain level of language proficiency is needed before a learner can use strategies s/he already has in L1 for L2:

Considerable support was found for the modified second hypothesis, namely that some sort of threshold or language competence ceiling has to be attained before existing abilities in the first language can begin to transfer. (20)

As for what is the nature of language problems that confront readers, he implies that it is largely lexical problems rather than syntactic problems:

More plausible, but elusive, is the notion that problems in foreign language reading which are due to language have to do with semantic and discourse processing, and are related to problems of conceptualization and, to put it crudely, word meaning. (ibid.)

Laufer's (1997) explanation might answer the question of why insufficient vocabulary knowledge hinders successful guessing:

Since the amount of information that can be cognitively manipulated at one point of time by controlled processing is limited, focusing on slightly or completely unfamiliar words will take up some cognitive capacity that would otherwise be used for higher-level processing of the text. (22-23)

Laufer (op. cit: 23) claims that the threshold vocabulary size is 3000 word families based on the result of her own research.

There seems to be enough evidence to believe that learners with vocabulary knowledge below a certain level cannot use inferencing strategies efficiently. Therefore, it might be necessary to change the focus of instruction below and above the threshold level. Below the threshold level, various activities focusing on building up vocabulary, perhaps using texts containing controlled vocabulary, would be beneficial. Once the learners reach the threshold level, more meaning-focused reading activities, using authentic materials,

with instructions of inferencing strategies, would be desirable.

3. Vocabulary Learning through Reading

In the previous section, we have seen that vocabulary is a major factor affecting reading comprehension in L2, and that a learner has to acquire considerable amount of vocabulary knowledge and appropriate strategies to infer the meaning of unknown words before s/he can reconstruct the writer's message satisfactorily. However, it is also widely recognized that reading itself can contribute to vocabulary learning. Schmitt (2000: 150) says: "There is a plenty of evidence that learners can acquire vocabulary from reading only", and cites the result of Saragi, Nation and Meister's (1978) experiment where learners learned the slang words of Russian origin just by reading a novel in which those words were used.

It seems plausible that vocabulary learning through reading is largely incidental in nature, but there are some researchers who claim that explicit instruction can enhance the process of vocabulary learning. In this section, I would like to look at several aspects of incidental vocabulary learning process first, and then consider the effectiveness of explicit instruction in reading activities for vocabulary learning.

3-1. Incidental Vocabulary Learning

As Schmitt (2000: 149) claims, the key to incidental vocabulary learning is to provide learners with as much exposure as possible. As Schmitt (ibid: 150) notes, incidental vocabulary learning can happen through exposure to spoken

discourse alone. However, as he (ibid.) claims, reading is more suitable exposure for vocabulary expansion than spoken discourse, at least for learners above a certain level, because spoken language tends to include only a limited number of low-frequency words.

The empirical research of Day, Omura and Hiramatsu (1991) showed that the treatment group who read a short story containing target vocabularies with ample frequencies and sufficient context did significantly better than control group who did not read the story on the vocabulary test. Although they do not try to elucidate how many times unknown words should be repeated in a text or how much (and what kind of) context should be provided for incidental vocabulary learning, the result of their study signifies that these two factors -- the number of repetition and quantity (and perhaps, quality) of context clues -- might be crucial. They (op. cit: 545) admit that exposure to unknown words only in one short story would not guarantee that they enter the long-term memory of learners, and acknowledge the need for repetitive exposure over time in different contexts.

The research of Day et al. indicates that incidental vocabulary learning can actually happen through reading, and implies efficacy of extensive reading as a means of vocabulary expansion. However, it should be noted that they used a "controlled" text in which target words -- words which researchers meant to be learned by learners -- were repeatedly used and sufficient contexts were provided for those words. Therefore, it seems questionable if it can be called 'genuinely'

incidental learning. Nevertheless, it could be said that their research is valuable in that it has implication for wider use of 'implicit' vocabulary instruction in EFL classrooms.

3-1-1. Lexical Inference and Vocabulary Learning

In section 2-3 above, we have considered lexical inference strategies in the context of reading comprehension. Lexical inference strategies have also been studied in the context of vocabulary learning. As Nation and Coady (1988) note, foci of studies are different according to the context in which researchers are considering lexical inference strategies:

Studies on getting the meaning give their attention to the types of clues available in context, learners' success or failure in using available clues, and the effort of training on using clues.

Studies on learning words from context sometimes consider the presence of clues, but are most interested in what has been remembered of a word from meeting it in context. (102-103)

However, it could be said that in order for the meaning of a word to be learned, it must be guessed from context in the first place, and that the same guessing strategies can be used as the ones we have seen in the preceding section. Therefore, what we want to know is what sorts of factors facilitate or hinder the retention of the meaning of a word after it has been guessed (hopefully, correctly) from context.

Schmitt (2000: 155) claims that the 'depth of mental processing' required for guessing the meaning of an unknown word is related to the extent to which it is retained in long-term memory afterwards: "... if the clues are sparser and the guessing process requires more cognitive effort, then the word is more

likely to be remembered.” He (ibid.) continues to state that in order to attain adequately deep level of mental processing the words must not be too easy to guess: “Thus guessability and retainability may have an inverse relationship.”

3-1-2. Incremental Nature of Vocabulary Learning

As we have seen in section 2-2, ‘knowing a word’ involves several aspects, and it is likely that each aspect develops gradually in different manners. Schmitt (2000: 117-118) states that only some sense of word form and meaning is likely to be picked up at the first exposure, and it is late in the acquisition process that such aspects as frequency, register and collocation are developed because knowledge about these aspects needs a large number of examples before it is fully acquired.

Schmitt (op. cit: 119) also acknowledges that distinction between receptive knowledge and productive knowledge is related to incremental nature of vocabulary learning. As he notes, traditional view is that a word is learned receptively first, and come to be used productively as vocabulary knowledge of the word fully develops. However, Schmitt claims that it is possible that a learner starts to use a word productively in a limited way before s/he masters its receptive aspects in full.

It is obvious that a learner has to encounter a word in different contexts over and over again before s/he fully develops vocabulary knowledge of that word. There seems to be no consensus about the number of encounters needed for

complete mastery of the word knowledge (probably because of such other factors as quality of attention, motivation etc.), but Nation (1990: 44) cites claims of several researches that at least 5-16 encounters are necessary. If that is the case, material writers will have to take great care so that important words are 'recycled' in various contexts in order to guarantee learners chance to learn different aspects of the word over time.

3-2. Can Explicit Instruction Play a Role?

As Day et al. (1991: 541) notes, it seems that there have been more and more researches supporting usefulness of incidental/implicit learning of vocabulary through reading in recent years. It might be the result of the shift from form-centered approach to meaning / communication-centered approach in language teaching in general. However, incidental / implicit approach seems to have a couple of flaws: 1) It is rather inefficient in that it takes considerable amount of time before a new word is fully acquired. 2) It is unpredictable which words are actually learned by learners. Therefore, it is desirable if some kind of intervention can be used to make up for these shortcomings.

Paribakht and Wesche (1997) address this issue. The result of their research indicates that learners who had vocabulary exercises as well as reading session learned target words better than learners who had reading session alone. They ascribe the gain to the nature of vocabulary exercises which were designed to make the target words salient in the text, and to require deep cognitive processing.

The empirical research of Paribakht and Wesche shows that explicit instruction can be incorporated into incidental vocabulary learning program to obtain the best result.

4. Summary and Conclusion

Reading and vocabulary learning are deeply interrelated with each other. We have seen that a certain level of vocabulary knowledge is necessary not only for reading but also for further learning of vocabulary. The existence of 'threshold level' implied by a number of researches seems to have an important implication for the pedagogy of vocabulary and reading. Priority should be given to building up the most basic vocabulary for beginners. For learners who have reached the level where they can read simplified texts, extensive reading using graded readers might be beneficial. More advanced learners might benefit from reading authentic materials with strategy instruction.

Incidental learning and explicit instruction have their own advantages and disadvantages, but they can be combined to make up for each other's weakness.

Bibliography

- Alderson, J. C. 1984. Reading in a foreign language: a reading problem or a language problem? In J. C. Alderson and A. H. Urquhart (eds.) *Reading in a Foreign Language*. New York: Longman, 1-27.
- Day, R., C. Omura and M. Hiramatsu. 1991. Incidental Vocabulary Learning and Reading. *Reading in a Foreign Language*, 7: 541-553.
- Hosenfeld, C. 1984. Case studies of ninth grade readers. In J. C. Alderson, and A. H. Urquhart (eds.) *Reading in a Foreign Language*. London: Longman, 231-249.
- Huckin, T. and J. Bloch. 1993. Strategies for inferring word-meanings in context: a cognitive model. In T. Huckin, M. Haynes and J. Coady (eds.) *Second Language Reading and Vocabulary Learning*. Norwood, NJ: Ablex, 153-176.
- Laufer, B. 1992. How much lexis is necessary for reading comprehension? In Bejoint, H. and Arnaud, P. (eds.) *Vocabulary and Applied Linguistics*. London: McMillan, 129-132.
- Laufer, B. 1997. The lexical plight in second language reading: Words you don't know, words you think you know, and words you can't guess. In J. Coady and T. Huckin (eds.), *Second Language Vocabulary Acquisition: A Rationale for Pedagogy*, pp.20-34. New York: Cambridge University Press.
- Nation, I. S. P. 1990. *Teaching and Learning Vocabulary*. New York: Newbury House Publishers.
- Nation, I. S. P. and Coady, J. 1988. Vocabulary and reading,

- In R. Carter and M. McCarthy (eds.) *Vocabulary and Language Teaching*. London: Longman, 97-110.
- Noro, T. 2002. The roles of depth and breadth of vocabulary knowledge in reading comprehension in EFL. *Annual Review of English Language Education in Japan*. 13: 71-80.
- Paribakht, T. S. and M. Wesche. 1999. Reading and "incidental" L2 vocabulary acquisition: An introspective study of lexical inferencing. *Studies in Second Language Acquisition*. 21: 195-224.
- Quian, D. D. 1999. Assessing the roles of depth and breadth of vocabulary knowledge in reading comprehension. *Canadian Modern Language Review*, 56: 282-307.
- Read, J. 1993. The development of a new measure of L2 vocabulary knowledge. *Language Testing*, 10: 355-371.
- Schmitt, N. 2000. *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.