Content Area Textbooks as Sources for Vocabulary Learning

Bambang Yudi Cahyono

Abstract: This study examines the potential of content area textbooks as a lexical environment for incidental vocabulary learning. The two questions proposed are how many rarely-used words content area textbooks contain within a particular number of words, and how different or similar the frequency of rarely-used words is in the linguistic studies and non-linguistic studies textbooks. This study concludes that content area textbooks contain a relatively high frequency of words within *Nat3* and *Nat4* levels. The high percentage of rarely-used words indicates that content area textbooks are rich sources for vocabulary learning. The findings also suggest that different kinds of textbooks in the content courses can be rich lexical environments for the study of second language vocabulary.

Keywords: vocabulary learning, lexical environment, incidental learning, content area textbook.

This article examines the potential of content area textbooks as sources for vocabulary learning. The underlying assumption of this article is that the large number of rarely-used words in textbooks of content areas would indicate a rich source of vocabulary learning for language learners, whereas the large number of frequently-used words would indicate that the vocabulary of content area textbooks was poor. It is hypothesized that the

Bambang Yudi Cahyono is a lecturer at the Faculty of Language and Art Education (FPBS), the State University of Malang (Universitas Negeri Malang).

content area textbooks used by Indonesian learners of English could be rich sources of vocabulary for learning English as a foreign language.

It is believed that developing knowledge about unfamiliar words can be facilitated through incidental learning from context. Encouraged by this proposition, Nagy, Herman, and Anderson (1985) conducted a study on how reading natural text influences vocabulary acquisition. Seventy eighth-grade students were involved in this study. The students were randomly assigned to read either a spy narrative or an exposition on river systems. Two measures of word knowledge, i.e. interviews and multiple choice questions, were specifically designed to detect the incidental learning of word meanings hypothesized to take place in contexts found in natural text. The students were then asked to do one of the versions of the vocabulary tasks. The results indicated that a greater proportion of the target words from a given passage were known by the subjects who had read that passage than by those who had not. It can be concluded from the study that incidental learning takes place while learners read.

Dupuy and Krashen (1993) carried out a study to investigate incidental vocabulary acquisition in French as a foreign language. The subjects were forty-two undergraduate students enrolled in French courses. Three classes were used, i.e. two intermediate French classes (one experimental and one control) and one advanced class (control). The students from the experimental class saw the first five scenes of the film *Trois hommes et un couffin* (without subtitles) to provide some background knowledge. The next day they were asked to read the script of the next five scenes in class. The two control groups did not see the film or read the passage. Then they were given a similar vocabulary test containing colloquial words which were not familiar to intermediate students of French. The results of this study indicated that the experimental group reached higher mean than either of the comparison groups. The research also found that the experimental group outperformed both control groups, confirming the proposition that vocabulary can be acquired incidentally by foreign language students.

Laufer (1992) investigated the relationship between passive vocabulary size and the comprehension of academic texts. Ninety-two subjects took part in the study. Two sets of reading comprehension tests and vocabulary size tests were assigned to the subjects to gain a reading score and a vocabulary size score respectively. The results indicated a

significant correlation between reading and vocabulary. Although the correlation did not imply a causal relationship, it was believed that vocabulary size is likely to be a predictor of reading comprehension. Furthermore, Laufer provided a practical implication by demonstrating how reading develops above the threshold level of 3000 word families. She argues that if the optimal reading level is, for instance, 70 per cent, then the vocabulary size to aim at will be 5000 word families; whereas if the passing grade is 63 per cent, then the vocabulary size needed is 4000.

The threshold level should be interpreted as justifying the idea that reading is a good way to learn words after the basic 3000 words are acquired. To learn these first words, reading might not be the most efficient way. If this is so, Laufer seems to add some perspective to what Krashen (1989) has found. Krashen suggested that the best way for second-language learners to develop their vocabulary in the second language is through reading. Additionally, the problem in understanding the relationship between reading comprehension and vocabulary size is the exact meaning of 70 per cent in terms of reading comprehension. If the questions in the reading comprehension were hard, 70 per cent would be good. If they were easy, it would be bad.

From the discussion of the issues in learning vocabulary it can be concluded that incidental learning of words can be facilitated through reading. In addition, in order to result in effective growth, a chance to learn words from reading should be provided for learners with a greater vocabulary size, that is, above the threshold level.

In addition to incidental learning from reading texts, the concept of lexical environments is also crucial for vocabulary learning. An example of lexical environment that has been studied is radio. Meara (1993) studied several sets of radio programs broadcast by BBC English, the teaching arm of the BBC's World Service. The programs were intended for learners of English at a variety of proficiency levels, that is from level 1, *beginner*, to level 9, *mastery*. Meara's study was to investigate the number and the difficulty level of the words in each text. The word list produced by Nation (1986) was used to judge the difficulty level of words in the programs. Words were classified from the least difficult, that is *Nat0*, up to the most difficult, *Nat4*. The analysis indicated that it is difficult to show that there was any straightforward relationship between lexical profiles of BBC English programs and the levels of learners they are aimed

at. In addition, the findings also suggested that all the radio programs studied were at about the same level lexically. From his study, Meara suggests the importance of knowing the lexical richness of other environments such as the classroom.

In responding to the lack of knowledge about classrooms as lexical environments as pointed out by Meara (1993), Meara, Lightbown, and Halter (1997) examined the availability of vocabulary in foreign-language classrooms. Their study was also based on the impact that communicative language teaching has brought to language teaching practices. In the perspective of the language teaching approach, teachers are unlikely to teach words explicitly or to assign students to memorize vocabulary lists. Transcripts of language classrooms and frequency of the words used by the teachers were analyzed according to lists developed by Nation (1986). It was concluded that classrooms do not seem to be rich lexical environments. However, it is suggested that the analysis might underestimate the true lexical richness of the classroom environments for the learners. At the end of the report, the researchers expect to be able to return to the question of the students' learning of the words which are available in classrooms. A remaining question is whether all teachers are more or less the same in providing linguistic input for the learners in the classroom.

Lightbown, Meara, and Halter (1999), then, compared three classroom lexical environments: audio-lingual classes, intensive classes taught communicatively, and adult classes taught communicatively with context-appropriate repetition. Audio-lingual classes were based on the principle that vocabulary should be limited in order to allow the learner to focus on linguistic features in sentences which were being imitated in language lessons. On the other hand, communicative classes emphasized vocabulary by reliance on target language words, although knowledge of grammar was still limited. Samples of equal length (in terms of the number of running words) from transcripts of teacher-student interaction were analyzed to measure lexical richness. Lemmas were divided into five bands based on the word lists created by Nation (1986). It was found that, in a variety of language teaching approaches, beginners and low intermediate learners were exposed to a very limited variety of new words. While the teachers in the intensive communicative classes had a greater variety of words in their transcripts, in general, the findings suggested that all three types of classrooms had a relatively impoverished vocabulary. The as-

sumption made by the authors, that classrooms provide access to new, unusual words in contexts that make them easy to learn, is not justified.

In the context of research which studies incidental learning and richness of lexical environments, the present study is aimed at examining the potential of content area textbooks as a lexical environment for incidental vocabulary learning. Indonesian learners of English as a foreign language in the English department of State University of Malang, Indonesia are the target of this study. In the first semester, the students are required to take an intensive course. In the next semesters, they have to take two types of courses, i.e. skill courses and content area courses. Skill courses that are aimed at providing skills in using English for communicative purposes include *Listening, Speaking, Reading,* and *Writing*. Content area courses include courses other than skill courses. These courses are intended to provide knowledge about language, literary studies, and teaching methodologies. The present study is focused on the potential of the content area courses for the learning of new words.

The data from the excerpts of content area textbooks will help answer two questions as follows: How many rarely-used words do content-course textbooks contain within a particular number of words? How different or similar is the frequency of rarely-used words in the linguistic studies and non-linguistic studies textbooks?

METHOD

In talking about managing words, two points must be agreed upon, i.e. the notion of word and the classification of the level of difficulty of words. According to Nation (1986), the concept of word is taken to be a word family. The knowledge of *develop*, for instance, subsumes the knowledge of *development, developmental,* and all their inflections. On the other hand, different meanings of the same form, e.g. *pupil*, would be different words. Nation's (1986) system to judge the level of difficulty of words is considered worth adopting (Meara, 1993; Meara, Lightbown, and Halter, 1997).

The word list produced by Nation (1986) was used to categorize the sample words. Following Nation's procedure, words were classified into five categories, that is *Nat0, Nat1, Nat2, Nat3,* and *Nat4*. The categories range from the least to the most frequently-used words. *Nat0* is the level in which functional words is categorized, for example, *he, do,* and *here*.

Nat1 is the first 1000 words beyond the basic lexical words and *Nat2* is the level in which the second 1000 words are categorized. Words of the *University Word List (UWL)* are those within the range of the 2000-3000 most frequently-used words, falling into the category of *Nat3*. The words beyond the first four categories are classified as words of *Nat4*. Judgment on the richness of the vocabulary in the textbooks is based on the use of words in those books belonging to or higher than the UWL level. This is because the UWL contains words required for study in the university level.

In each course, at least one required textbook is commonly used. Only one required textbook was taken from each course to be the sample textbook. Word samples were then taken from ten different content area textbooks. The sample courses used in this study are categorized into two, i.e. linguistic studies courses and non-linguistic studies courses. The five courses in Linguistic studies are *English phonology* (Grimson, 1989:16), *morphology and syntax* (Francis, 1958:164), *introduction to linguistics* (O'Grady and Dobrovolsky, 1992:374), *discourse analysis* (Stubbs, 1982:15), and *sociolinguistics* (Chaika, 1982:44). The non-linguistic studies include *teaching English as a foreign language* (Harmer, 1991:16), *language testing* (Oller, 1992:38), *teaching English for young learners* (Halliwel, 1992:12), *language teaching media* (Richard-Amato, 1996:169), and *introduction to research methods* (Nunan, 1992:18). From each textbook 120 words were taken as the word samples. The sample words are words from excerpts taken randomly from the textbook. The complete excerpts can be seen in Appendix 1.

RESULTS

The analysis of the lemma types in the content area textbooks resulted in two kinds of findings. The first is the distribution of lemma types in the content area textbooks. The central tendency of the distribution in terms of the average of the lemmas from the content area textbooks was also observed. The second is a comparison of lemmas of the linguistic studies and the non-linguistic studies textbooks.

Distribution of lemma types in the content area textbooks can be seen in Figure 1. In this figure, frequency (f) and percentage (%) of the lemmas are presented from the least (*Nat0*) to the most (*Nat4*) frequently-used lemmas.

Figure 1 Distribution of lemma types in the content area textbooks

Category	Content areas		Nat0	Nat1	Nat2	Nat3	Nat4	Σ
	Textbooks							
Linguistics	English phonology (EP)	f	20	31	8	11	4	74
		%	27.03	41.89	10.81	14.86	5.41	100
	Morphology and syntax (MS)	f	19	35	11	2	2	69
		%	27.54	50.72	15.94	2.90	2.90	100
	Introduction to linguistics (IL)	f	15	24	5	4	4	52
		%	28.85	46.15	9.62	7.69	7.69	100
	Discourse analysis (DA)	f	20	24	16	13	8	81
		%	24.70	29.63	19.75	16.05	9.87	100
	Sociolinguistics (SL)	f	20	28	7	10	8	73
		%	27.40	38.36	9.59	13.70	10.95	100
Non Linguistics	Teaching English as a foreign language (TEFL)	f	16	33	4	8	3	64
		%	25.00	51.56	6.25	12.50	4.69	100
	Language testing (LT)	f	10	37	3	11	5	66
		%	15.15	56.06	4.55	16.67	7.57	100
	Teaching English for young learners (TEYL)	f	17	35	4	6	2	64
		%	26.56	54.69	6.25	9.37	3.13	100
	Language teaching media (LTM)	f	16	42	2	10	11	81
		%	19.75	51.85	2.47	12.35	13.58	100
	Introduction to research method (IRM)	f	18	42	3	8	1	72
		%	25.00	58.33	4.17	11.11	1.39	100

From the raw data presented in Table 1, the distribution of lemma types are presented graphically in Figure 2. This figure shows the number of different lemmas from ten different content-course textbooks.

Figure 2 Distribution of lemma types

Figure 2 Distribution of lemma types
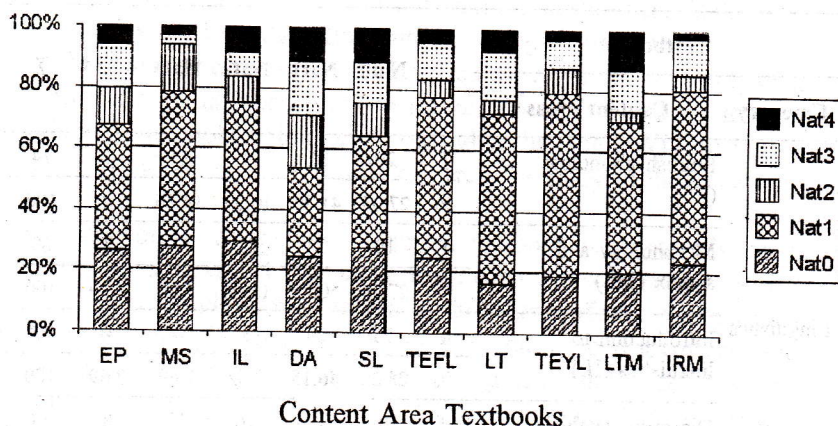
DISCUSSION

There were, on average, 69.6 lemma types from the ten content area textbooks combined. It was found that about one-fourth of lemmas come from *Nat0* (24.57%), almost half of the lemmas come from *Nat1* (47.56%), and about one-tenth of the lemmas come from *Nat2* (9.05%). The number of lemma types within *Nat3* and *Nat4* is 11.95%, and 6.87%, respectively. In addition to the distribution of lemma types in all ten different textbooks, a comparison was made to examine the lemmas of the linguistic studies and the non-linguistic studies textbooks.

In the linguistic studies textbooks, there were on average 69.8 lemmas. It was found that more than one-fourth of lemmas come from *Nat0* (26.93%), about two-fifths of the lemmas come from *Nat1* (40.69%), and about one-seventh of the lemmas come from *Nat2* (13.47%). The number of lemma types within *Nat3* is 11.46% and within *Nat4* is 7.45%. Both levels together comprise about one-fifth of the total number of the lemma types (18.91%).

In the non-linguistics studies textbooks, there were 71.4 lemma types. It was found that about one-fourth of the lemmas come from *Nat0* (24.37%), more than half of the lemmas come from *Nat1* (52.94%), and about one-twentieth of the lemmas come from *Nat2* (12.5%). The number of

lemma types from *Nat3* is 12.04% and from *Nat4* is 6.21%. The two levels make up about one-fifth of the total number of lemma types (18.25%).

It is interesting to observe the results of comparison of the numbers and the percentages. Both categories of content area textbooks have approximately similar average number of lemmas, i.e. 69.8 in the linguistic studies textbooks and 71.4 in the non-linguistic studies textbooks. The relatively similar number of lemmas from the same size of word samples (120 for each textbook) may indicate that the two categories of textbooks are not different in the richness of lexical input. This is also evident in the finding that the two categories of textbooks also have relatively similar percentage of words within *Nat3* and *Nat4*, i.e. 18.91% and 18.25% for the linguistic and non-linguistic studies textbooks, respectively.

CONCLUSION AND SUGGESTION

It may be concluded from this study that content area textbooks contain a relatively high frequency of words within the *Nat3* and *Nat4* levels. The high percentage of rarely used words, about one-fifth of the total number of the sample words for both linguistic studies and non-linguistic studies textbooks, indicates the potential of content area textbooks as sources for vocabulary learning. The findings also suggest that different kinds of textbooks in the content courses can be rich lexical environments for the study of second language vocabulary.

REFERENCES

Chaika, E. 1982. *Language the Social Mirror.* Cambridge: Newbury House.

Dupuy, B., and Krashen, S.D. 1993. Incidental Vocabulary Acquisition in French as A Foreign Language. *Applied Language Learning, 4,* 55-63.

Francis, W.N. 1958. *The Structure of American English.* New York: The Ronald Press.

Grimson, A.C. 1989. *An Introduction to the Pronunciation of English.* London: Edward Arnold.

Halliwel, S. 1992. *Teaching English in the Primary Classroom.* New York: Longman.

Harley, B., Howard, J., and Rogerge, B. 1996. Teaching Vocabulary: An Exploratory Study of Direct Techniques. *The Canadian Modern Language Review, 53,* 281-304.

Harmer, J. 1991. *The Practice of English Language Teaching*. New York: Longman.

Krashen, S.D. 1989. We Acquire Vocabulary and Spelling by Reading: Additional Evidence for the Input Hypothesis. *Modern Language Journal,* 73, 440-464.

Laufer, B. 1992. How Much Lexis is Necessary for Reading Comprehension? In Arnaud, P. & Bejoint, H. (Eds.) *Vocabulary and Applied Linguistics*. London: Macmillan.

Lightbown, P.M., Meara, P., & Halter, R.H. 1999. Contrasting Patterns in Classroom Lexical Environments. In Albrechtsen, D., Henriksen, B. & Mees, I. (Eds.). *Perspectives on Foreign and Second Language Education: Festschrift in Honour of Kirsten Haastrup*. Odense, Denmark: Odense University Press.

Meara, P. 1993. Tintin and the World Service. *IATFL Annual Conference Report,* 32-37.

Meara, P., Lightbown, P.M., & Halter, R.H. 1997. Classrooms as Lexical Environments. *Language Teaching Research 1,* 28-47.

Nagy, W.E., Herman, P.A., & Anderson, R.C. 1985. Learning Words from Context. *Reading Research Quarterly,* 20, 233-253.

Nation, I.S.P 1986. *Word Lists: Words, Affixes and Stems*. Wellington: Victoria University English Language Center.

Nunan, D. 1992. *Research Methods in Language Learning*. Cambridge: Cambridge University Press.

O'Grady, W., and Dobrovolsky. 1992. *Contemporary Linguistic Analysis*. Toronto: Copp Clark Pitman.

Oller, J.W. 1992. *Language Tests at School: A Pragmatic Approach*. London: Longman.

Richard-Amato, P.A. 1996. *Making It Happen*. New York: Longman.

Stubbs, M. 1982. *Discourse Analysis: The Sociolinguistic Analysis of Natural Language*. Oxford: Basil Blackwell.