English Department Students' Collocation Abilities

Siusana Kweldju
Universitas Negeri Malang

Abstract: *In* order to discover how well English Department students' knowledge of collocations and how successful the communication strategies students used to help them supply the expected collocations, 60 subjects were involved in this study. A fill-in-the-blank form test was administered to them to supply the collocates of the given nodes of selected collocations. The result of the analysis indicated that learners had a low mastery of collocations, and collocations need deliberate learning and teaching.

Key Words: collocation, vocabulary

This present paper addresses the issue of collocations as an important aspect in EFL instruction. Finding a right word in a second language is not easy, but through the literature of second language acquisition research, there were still a few researches done to investigate English language learners' knowledge of collocations, and there were even very few in Indonesia, if any. Although little was done to measure collocation capability, it could be predicted that Indonesian learners' knowledge on English collocations was still limited. The prediction above was made after reviewing a number of vocabulary size studies at different levels. Bahns & Eldaw (1993) have discovered that learners' knowledge of collocations lagged behind their vocabulary knowledge in general. Vocabulary size studies are only restricted to measure how many receptive or productive words students have, in terms of either the depth or the breadth of vocabulary. Collocation ability, however, does not only deal with knowing the meaning of words but also the combinability of multi-word units.

The results of vocabulary size studies revealed that English language learners in Indonesia were still lack of sufficient vocabulary. For example, the receptive vocabulary size of the seniors of 15 English Departments averaged about 4664 base words or 2800 word families only (Kweldju, 1997). The vocabulary size of the first year students of the non-English Departments in a university in Indonesia was only estimated to be 1226 words (Nurweni, 1995), and in one prestigious university was only about 2520 (Nuswantara, 1998). Concerning the productive vocabulary, the English Department students in two different English Departments in Indonesia had about 1800 productive words only (Hamdi, 1998; Abdulla, 1998).

In second language teaching vocabulary has been neglected. Therefore, upon leaving a language program students often find themselves at a literal loss of words in the natural environment, and how those words combine into fixed phrases and more close models. Vocabulary instruction should begin very early, and its learning demands perseverence.

The importance of collocations

Knowing a word is not only knowing its form and meaning. Words occur together to make collocational patterns, or they do not combine and recombine freely and randomly with others (Hanks, 1987). Using language is to select more than one word at a time, and to blend the selections with each other (Sinclair, 1987). This constraint takes place at different degrees to make a certain word follow certain others (Farghal & Obiedat, 1995). A number of studies in Russia discovered that L2 learners knew a little about these constraints and their errors constitute a high percentage of all errors committed (Biskup, 1994; Bahns & Eldaw, 1993). To compensate this, they inclined to paraphrase them as their communication strategy although collocations are not easy and even more difficult to paraphrase (Bahns & Eldaw, 1993), and to clutch for the words learners feel safe with or the lexical teddy bears (Hasselgren, 1994).

Some reasons why collocations were neglected

That we neglect collocations was probably much influenced by one of the fundamental assumptions of transformational grammar, that human mind had

the mental capabilities to use language in a free, creative and novel fashion. Thus, most sentences we encounter are novel, the term used in Chomsky's (1957) earlier work, and in all Chomskyan models a characteristic of competence is its creative aspect (Cook, 1988). Chomsky (1972) argues that this essential creativity of language indicates that language can't simply be learned through imitation, but it enables human beings to cope with sentences which have never been heard or produced before.

Kennedy (1990), for example, is one of those who does not quite agree with the exaggeration of language novelty, but also realized that Chomsky's (1957) earlier work was made only for reacting to the behaviourist's operant conditioning and verbal chaining. Antal (1985) has highlighted that creativity is the property by the speaker, but not the language. If a language was always grammatically new, or lexically new, it may impose incomprehensibility and misunderstanding. Also, that a speaker is creative is only the extent of being able to combine the old elements within the bounds permitted by the old grammar in a novel fashion. Even Chomsky (1976) in his later work admits that novelty is constrained by the established knowledge in the mind.

Also, as long as perception is concerned learners find no difficulties, because collocations are fully transparent and comprehension tests produced 100% correct answers (Biskup, 1994).

Collocations are unique

Collocations are unique, and some are unpredictable. The same conceptual organization can be realised differently in different languages, and only a native speaker has a sufficient experience to produce the collocations of his language, but not a non-native speaker whose environment only gives him a weak trace or no trace at all for a later recall. A conscious effort and different mental operations are crucial to them.

Collocations in second language learning

When a lexical item like *fast* is introduced in the classroom, it is not enough to explain what it means. To enable a learner to use it in production, a learner

should also need to know that the word "fast" occurs both in open-slot and close-choice/stable collocations. *Conduct a research* is an example of the close choice principles. The failure to recognize the collocability of *conduct* with *research* is a direct consequence of learning and teaching words individually rather than collocationally. *To commit a murder* is a far more fixed collocation than those with other verbs, like *to investigate*, and *to witness a murder*. Collocations can be considered as memorized whole utterances or phrases. They are learned through rote memorization or overusing them (Krashen & Scarcella, 1978; Hakuta, 1974). Collocations underline the importance of teaching language in chunks or memorized wholes (Hasselgren, 1994), because many linguists have suggested that collocations, like idioms, cliches and non-canonical forms were stored as patterns.

Knowledge of collocations enables a learner to say *I'd like a peanut butter and jelly sandwich* in American English rather than *I'd like a jelly and peanut butter sandwich* (Oxford & Scarcella, 1994). A native speaker has a thorough intuition of these possible combinations. For example, *broad* is synonymous with *wide, large and extensive*. It is used not only with physical entities such as *roads and rooms*, but also with more abstract notions, like *broad education* and *broad mind*. However, the word *large* cannot go together with *road, education* and *mind*. And although both *wide* and *broad* can go together with *street* and *river*, only *wide* can go together with *open*, like (1) but not *broad*, like (2):

(15)   He opened the door wide.

(16)   *He opened the door broad.

Similarly, we can say *an extensive view*, but we cannot say *an extensive river*, as presented in Table 1:

Table 1 Nouns-Adjective Collocations with *broad, wide and extensive*

Adjective	Noun							
broad	river	street	mind	field	-	-	-	-
wide	river	street	-	field	network	experience	-	-
extensive	-	-	-	-	network	experience	view	research

The effects of using wrong collocations

Second language learners do not have sufficient knowledge on collocations and they normally find refuge to lexical teddy bears, which are normally common words with wide ranges of meanings. They are learned in the early stages and are systematically overgeneralised by advanced learners. The effect of using lexical teddy bears occupies the zone of using between right and wrong words. The wrong lexical choice may cause a mismatch, distorted message, the use of non-existent word or phrase, highly loaded words, and collocational dissonance. This may cause disharmony without being deviant in meaning or style. For example, the use of words which does not only cover its own area but spreads over the terrain of some other words, like the use of *admit a discount*, instead of *allow or grant a discount* (Hasselgren, 1994).

Second language learners normally had favourite preferences of lexical teddy bears. Usually they were general and neutral words, like the use of intensifiers *very* (much), *a lot* (of) and *extreme*(ly). Native speakers, however, tended to choose intensifiers which cover specific areas only. Norwegian learners, for example, are much more likely than native speakers to use the three intensifiers above for any contexts.

Purpose

This study was conducted to investigate English Department students' collocation abilities, whether they still needed some explicit learning in collocations or not:

1. Do students have sufficient knowledge on collocations?
2. Can their communication strategies help them supply the expected collocations?

METHOD

Subjects

There were 60 subjects involved in this study. They were Indonesian native speakers who were studying as the fifth semester students at a state teachers' college in Indonesia and who were preparing themselves to teach English as a foreign language to other Indonesian native speakers in high schools. They had received an average of ten years' instruction in English.

Instrument

Collocations exist in any languages, including English. There are two kinds of collocations: grammatical and lexical. This study investigated lexical collocations only. An instrument, a fill-in the blank form, was administered for eliciting the data. This task was to test the subjects' production ability in providing the right words which collocate with some given words. There were 25 items in this fill-in-the-blank form. Items numbers 1 through 15 were developed by Bahns and Eldaw (1993), and 16 through 25 by Farghal and Ibiedat (1995).

In items numbers 1 through 15 the subjects' knowledge to investigate was the verb+noun collocations with the verbs as the collocates and the nouns as the nodes. Verb+Noun collocations are relatively common, and a frequent source of difficulty for EFL learners (Benson, 1985). The collocations were put in sentences, and were tried out to 2 native speakers, who provided the same collocations for the given nodes (Bahns and Eldaw, 1993).

Items numbers 16 to 25 were developed based on 22 common English collocations of topics such as food, clothes, and weather. Two native speakers were also involved for the validity of the collocations (Farghal and Ibiedat, 1995).

Before the task was administered, 26 items were tried out to 8 intermediate learners of English, whose TOEFL scores averaged 425. The results were consulted to two native speakers of English, a psychologist and a writer, and one item was determined to discard. A rivision was also made for the stem of one item, and more possible answers were added to the key, e.g. although *pay+compliment* sounds more educated and collocational for a native speaker,

give+compliment is also acceptable. Although *spicy+food* is more precise, *hot+food* is fine. Although *do+damage* is more collocational, *cause+damage* is also acceptable. Both *achieve* and *attain* can equally go with *perfection*, and although *bland soup* is more collocational, *plain soup* is acceptable.

Before the instrument was administered it was made sure that all subjects knew the wording used in the instrument.

FINDINGS AND DISCUSSIONS

Students had Limited Knowledge of Collocations

The results of this study supported Bahns and Eldaw's (1993) finding. The elicitable target collocations from each student ranged between 1 and 11 collocations or 4 to 44%, as shown in Table 1. Most of the students or 82% could only make 2-6 correct answers. In average students produced 4 collocations only or 16%.

As shown in Table 3 only 19 collocations were successfully elicited, and 6 were not elicited at all. If the more open class alternative was not taken into consideration, actually there were only 16 collocations to elicit. This is because students did not know the combinations of *do damage, pay compliments*, and *bland soup*, which are more collocational, and they substituted them with *cause damage, give compliments*, and *plain soup*, which are also acceptable.

Table 2 Percentage of Correct Collocations Made by Students

Category	Total C.	Collocations	f	%
Known (75 - 100%)	2	hot food (temperature)	45	75
		spicy (22)/ hot food (23)	45	75
Fairly known (50 - 74%)	1	do (0)/ cause (31) damage	31	52
Little known (25 - 49%)	4	lenient rules	26	43
		attend the lectures	19	32
		cancel the order	18	30
		with draw some money	18	30
Hardly known (1 - 24%)	12	rough/ stromy/ wild sea	12	20
		admit one'siswa defeat	11	18
		pay (0)/ give (11) compliments	11	18
		reject one'siswa proposal	7	12
		heavy drinker	6	10
		refuse admission	6	10
		weak tea	4	7
		achieve/ attain perfection	4	7
		Bland (0)/ plain (3) soup	3	5
		set one'siswa watch	2	3
		height of the summer	2	3
		take the call	1	2
Unknown	6	keep a diary	0	0
		serve the sentence	0	0
		arouse compassion	0	0
		whip the cream	0	0
		rich/ fatty food	0	0
		fast colour	0	0

Although students could use those 3 alternatives above, they could not automatically find refuge to the more open class alternatives. That *cause, give* and *plain* belong to the first 1000 most common words (Nation, 1994), students were assumed to know their meanings, but only 52% of them could use *cause*. Less subjects could use *give* (31%), and only 5% could use *plain*. Besides the unguessability of collocations, these differences might link to word frequencies; although *plain* belongs to the 1000 most frequent words, it is the least frequent

compared to the other two. The Cobuild Wordlist indicates that *give* was the most often used in the dictionary explanations (431 times); *cause* was used 249 times; and *plain* was unlisted.

Collocations Students Knew the Most and the Least

There was 1 collocation most known by students, i.e *hot food* in terms of both the height of the temperature or the burning taste. Students even preferred to use *hot food* than *spicy food* for the burning sensation, although *spicy food* is more

Table 3 Known and unknown target collocations and numbers and percentages of students to know them

Target collocations		Elicited Collocates
Collocate	node	
set	watch	fix, repair, correct, suit, adjust.
attend	lectures	meet, see, join, obey, listen
withdraw	money	cash money, liquidize money, clear, draw.
reject	proposal	cancel, refuse, return, stop, postpone, draw
refuse	admission	drop, miss, retire, reject, fire, expell
whip	cream	stir cream, blend cream, mix cream
achieve	perfection	finds, reach, gol, accomplish, gain, attain, acquire
lenient	rules	soft, weak, unstrictly, tender, flexible, permissive, rules

Code: Total C= total collocations

precise for a native speaker. That students were very successful to produce *hot food* seemed to be the product of excessive exposure and over-learning, or even both.

None supplied the expected *do damage*, but 52% of subjects produced *cause damage*. Actually both *do* and *cause* are high frequency words, but it was the latter which was produced by the students. The answer seems to be that the latter is a parallel of *menyebabkan kerusakan* in the subjects' native language, and it is more predictable as it matches the logical link of cause and effect. This indicates that some collocations must be deliberately learned, and some others are not equally needed.

A small number of students (30-43%) could still produce *withdraw +money, cancel+order, attend+lecture,* and *lenient+rule.* The first two collocations might be learned in ESP courses and the last two in their daily conversation as students.

Twelve collocations were hardly known. Among these 12 collocations, 2 target collocates were not produced as mentioned above. According to the native speaker, however, *pay compliments* was more collocational, and so was *plain soup,* because *plain* which means simple could go with *person, city, clothes, etc.,* but, *bland* was more common to go with food. Six collocations were completely unknown, i.e. *keep a diary, serve a sentence, arouse compassion, whip the cream, rich/fatty food,* and *fast colour.*

Lexical Teddy Bears Used as Compensating Collocates

One lexical teddy bear elicitable was *take. Take* was used 44 times to combine with different nodes, but, unfortunately, it was not used with the node *call,* as expected. Students used *take* to go with *lecture, money, sentence.* While *take+lecture* and *take+ sentence* are wrong, *take+money* is impossible in the given sentential context, although possible in *take+money from one's account.* This also indicates that collocations are unguessable.

Reliance on L1 as a Communication Strategy

Relying on L1 can either be a negative or positive transfer. Sometimes students overlooked that every language has its collocations, and those collocations are stable. Therefore, instead of producing *keep a diary,* 75% students produced *write a diary* which is very appropriate in their L1, but not in English, unless *write in a diary.* It is obvious that L1 was used for their communication strategy, and even the most frequently manifested. Other examples include *accept sentence* (10 cases) as a literal translation of *menerima hukuman* for *serve sentence,* i.e. *match+watch* for *set+watch* (11 cases), *follow+lecture* for *attend+lecture* (6), *take+money* for *withdraw+money* (24).

Using The Open Choice Principle and Giving Approximately Synony-mous Words or Contextual Words

Although students might transfer their L1 knowledge as shown above in solving the problem of their L2 production, this transfer strategy was not always used, even when they could benefit from it. For example, only 11 students produced *admit defeat*, which was parallel to his L1. Most of the elicited words were not based on the L1 transfer, i.e. *have defeat (3); show defeat (10), get defeat (5); accept defeat (3), confess defeat (1), express defeat (1),* and *face defeat (1)*; only 18 subjects supplied *cancel+order*, which was parallel with *membatalkan pesanan*, while the rest 41 gave 23 approximate answers like *delay, postpone, stop, refuse+order*; none knew *serve sentence*, and 30 subjects randomly used approximate words, like *undergo, go, sign, take, do, choose, stay, make+sentence, etc.*

The other open choice answers produced by subjects were as follows:

Table 4 Subjects' Approximate Collocates Elicited

Correct Answer		Student	
Total	%	f	%
1	4	2	3
2	8	9	15
3	12	19	32
4	16	10	17
5	20	5	8
6	24	6	10
7	28	4	7
8	32	3	5
9	36	1	2
11	44	1	2

Guessing Could not Help Collocate Production

Normally students did not want to use wild guessing. They preferred not to give any responses. This indicates that students were aware that guessing was difficult to make. Guessing was more possible for one word but not another. For example, that *light food* was antonymous with *rich food* was unpredictable. The target collocation *rich/fatty food* was unknown for students, except two who gave *fatty food*, but none supplied *rich food*. As induced by the instrument the antonym of *rich/fatty food* was *light food*, and based on this 26 out of 59 students gave

heavy food; 2 even gave *dark food*, and the rest gave some features of *heavy*, like *big food, strong food, hard food,* and *fat food*. Eleven students still refrained from using their guessing strategy.

Therefore, students guessed it to be *heavy food*. At another time when students could have guessed that the opposite of *strong tea* was *weak tea*, none guessed so, instead 21 subjects produced *light tea*. The subjecs might not know that *weak* could describe the quality of *liquid*. Although *strong* was not antonymous with *light* and the like at all, they preferred to use *light tea, soft tea, mild tea, fair tea,* and *plain tea* than *weak tea*. Still 9 subjects did not try to respond at all.

One of the opposites of *moderate* is *conservative*. But the opposite of *moderate drinker* is *heavy drinker*. Being unaware of homophones, 13 subjects produced *conservative drinker*, some others supplied *radical, traditional, extreme, conventional, strict, strong, maniac,* and *old drinker*. There were only 6 subjects who supplied the target collocation.

Similar Collocates Used in a Different Context

The problem became more complicated when there were similar pairs of collocations, but they were used in different contexts. In English either *answer* or *take* collocates with *call*. Since students knew the first but none knew the second, 43 out of 59 subjects overgeneralized *answer+call* to a situation in which the phone had been answered but had not been taken by the target receiver.

Another example was the use of *take+money* by 23 subjects. *Take* is possible to use in a context such as *take money from my account* but it is not equally acceptable for *withdraw money*.

The Intralingual Strategy

In English we normally say *the depth of winter*, students overgeneralized this to *summer*, and they gave *the depth of summer* for *the height of summer*. *Giving the Descriptive Features of the Node*

When the students were supposed to give the target collocation *height of summer*, instead of producing *height*, students resorted to the most likely features of summer, like *hot of summer, heat of summer, warmth of summer, dry of sum-mer.*

Another example is for the target collocation *rough/ stormy/wild sea*. Only 12 students knew the collocates, and 17 of them refrained from giving any answers, and the rest tried to use 10 different variants, which were all descriptive in nature, like *waving sea, hard sea, tidal sea, wavy sea, heavy sea, high sea, windy, cruel, strong,* and *furious*. All these alternates greatly varied and more dependent to the subjects' perception of a rough sea.

CONCLUSION

This study has shed light on the serious deficiency of the mastery of collocations, because students could only supply 16% of the expected collocations; 82% of subjects could only produce 2-6 collocations. Nine out of 25 collocations were not elicited at all.

Collocations should be brought to the attention of the learners and deliberately learned, because it was evident that excessive exposure and over-learning were crucial for retainment. First, a low frequency collocate like *withdraw* in *withdraw+money* was more likely to produce after having been deliberately learned than high frequency ones like *keep* in *keep+diary*, and *rich* in *rich+food*, etc. Second, limited knowledge in collocations led the learners to resort to lexical teddy bears in their guessing strategy, which ended up with wrong collocations. However, when the same lexical teddy bear was supposed to use for a target collocation, learners did not use it and made wrong guesses. This shows that collocations are stable as a genre of multi-word unit, but unpredictable. Third, reliance on L1 as a production strategy did not always result in positive transfer because the one-to-one correspondance hypothesis held in only few cases. Even when there was a one-to-one correspondence between L1 and L2 collocations, students could not confidently transfer their L1 knowledge for solving their L2 production problem. Fourth, students became more confused when the collocate of a collocation was a homophone. e.g. they did not know that *light food* was antonymous with *rich food*.

Students inclined to use the more open-class collacations than the close-class ones. *Hot food*, which belongs to the open-class were elicited much more often than *spicy food*, which belongs to the close class, although the latter was more precise. Open class collocations were easier to guess when there was a one-

to-one correspondence between L1 and L2 and the collocation matched a logical link, like cause and effect, i.e. *cause damage*.

The L1 transfer strategy was the most frequent manisfested in collocation production, but learners also used the intralingual transfer, and the descriptive feature of the node.

Without sufficient knowledge of collocational restrictions it is impossible to develop an idiomatic and natural skill in English, because not every word can enter open-slot collocations like *hot food*, and *cause damage*.

REFERENCES

Abdulla, R. 1997. *English Department Students' Vocabulary Competence in Relation to their Speaking Fluency of FKIP Syiah Kuala University Darussalam Banda Aceh.* An unpublished master's thesis: IKIP Malang.

Antal, L. 1985. Is it really true that nearly all sentences are novel ones? *Word*, 36(3).

Benson, M. 1985. A combinatory dictionary of English. *Dictionary*, 7.

Bahns, J. & M. Eldaw, 1993. Should we teach EFL students collocations? *Systems* 21(1).

Biskup, D. 1992. L1 influence on learners' renderings of English collocations: a Polish/ German empirical study. In P.J.L. Arnaud & H. Bejoint, *Vocabulary and Applied Linguistics*. London: Macmillan Academic and Professional Ltd.

Chomsky, N. 1957. *Syntactic Structures*. The Hague: Mouton.

Chomsky, N. 1972. *Language and Mind*. New York: Harcourt Brace Jovanovich.

Chomsky, N. 1976. *Reflections on Language*. London: Temple Smith.

Cook, V.J. 1988. *Chomsky's Universal Grammar*. Oxford: Basil Blackwell Ltd.

Hakuta, K. 1974. Prefabricated patterns and the emergence of structure in second language acquisition. *Language Learning*, 24(2), 287-297.

Hamdi, M.Z., 1998. *Productive Vocabulary Mastery in Composition Ability of the Students of STKIP Hamzanwadi Selong*. An unpulished Master's Thesis: IKIP Malang.

Hanks, P. 1987. Definitions and Explanations. In Sinclair, M. *Looking UP*. London, Collins ELT.

Hasselgren, A. 1994. Lexical teddy bears and advanced learners: a study into the ways Norwegian students cope with English vocabulary. *International Journal of AppliedLinguistics*, 4(2).

Hornby, A.S. 1975. *Guide to Patterns and Usage in English.*London: The English Language Book Society & Oxford University Press.

Kennedy, G.D. 1990. Collocations: where grammar and vocabulary teaching meet. In S. Annan (Ed.) *Language Teaching Methodology in the Nineties*. Singapore: RELC.

Krashen S. & R. Scarcella, 1978. Routines and patterns in language acquisition and performance. *Language Learning*, 28(2).

Kweldju, S. 1997. *English Department Students' Vocabulary Size and the Development of a Model of Extensive Reading with Individualized Vocabulary Learning*. Singapore: SEAMEO-Regional Language Centre.

Nation, P. 1994. *Vocabulary lists: words, affixes and stems.* English Language Institute, Victoria University of Wellington, New Zealand.

Nurweni, A. 1995. *The English Vocabulary Knowledge of the first year students in an Indonesian University Students*. Unpublished MA Thesis. Victoria University of Wellington.

Nuswantara, K. 1998. *The Effectiveness of Vocabulary Guessing Strategy to Read Content Textbooks Used by Students with Different Vocabulary Sizes*. An unpublished master thesis:IKIP Malang.

Sinclair, J. 1987. Grammar in the dictionary. In J. Sinclair, *Looking Up*. London: Collins ELT.