

# AN ANALYSIS OF COLLOCATIONAL VIOLATIONS IN TRANSLATION

---

■ Sirinna Boonyasaquan

## Abstract

*This study aims to raise awareness in teaching collocations to EFL/ESL students among teachers, curriculum writers, and related personnel, in the belief that collocations play a crucial role in learning the English language. The author presents the results of her research on collocational violations in translation by demonstrating the patterns of the violations and attempting to give plausible accounts for the errors so as to help avoid problems in learning a foreign/second language and promote effective communication.*

## INTRODUCTION

Vocabulary is most essential in communicating in a foreign/second language and it is even regarded as more important than grammar. A learner who is very bad at grammar may not be able to convey the precise message; however, if s/he does not know the exact word or phrase to communicate, s/he may not be able to convey any information at all (Farghal and Obiedat, 1995). However, when learning a new word, it is important to learn it with its frequent co-occurrence/s or word partner/s, or what is

called a collocation. Learning a new word in isolation is not meaningful and can easily cause problems for learners.

In Thailand, collocations seem to be neglected in classroom teaching. Very few teachers raise students' awareness of collocations or stress that English collocation exists and being able to use correct collocations greatly helps a learner to master the language. James (1998) has asserted this point that "adherence to the collocational conventions of a foreign language contributes greatly to one's idiomacity

and nativelikeness, and not doing so announces one's foreignness" (p.152). Herbst (1996) confirms this by saying that "...competence in a language involves knowledge about collocation" (p. 389). Thai teachers appreciate very little the significance of collocations and action research on this topic is rarely undertaken although students' collocation errors are frequently observed. On the contrary, in Taiwan, Japan, China, Korea, etc., Thailand's Asian neighbor countries, collocation is a very popular research topic (Huang, 2001; Kumiko & Masakazu, 2001; Park & Kakehi, 2001; Hsu 2002; Yu & Li, 2003).

Collocations should be introduced to students at the earliest possible stage of language learning. It is not an easy task for EFL/ESL learners to acquire collocations especially if they are not properly taught. Obviously, the sooner collocations are introduced to a learner, the better chance for her/him to be successful in learning the language. In order to be successful in teaching and learning collocations, those concerned should realize that collocation is an important aspect of language learning, in the same way as grammar, pronunciation, stress, and other language components (Hill, 2000).

The results of this study have revealed fruitful information about the violations of collocations made by Thai students in general. Plausible explanations and insights into solving and avoiding collocational problems, resulting

in effective communication, have been attempted. Awareness of collocations among EFL people in Thailand have been raised and constructive measures for bringing collocations into classrooms have been suggested as a matter of urgency.

In the first place, the study set out to examine students' ability in translating a business text from Thai into English, especially with regard to their application of collocations in their translated work. Next, the patterns of collocational violations that the students produced in their translations were critically investigated. Third, a plausible explanatory account for the collocational violations was attempted.

## LITERATURE REVIEW

### What is collocation?

Most linguists define collocation in a similar way. Baker (1992) has given a definition of collocation as the tendency of certain words to co-occur regularly in a language. She has indicated that words are not tied together at random in any language; on the other hand, there are always restrictions on the way they can be combined to convey meaning. Baker (1992), and some linguists, such as Bahns (1993) and Benson (1985), have proposed two major types of collocation: lexical and grammatical. Lexical collocations are combinations of open class words such as verbs, nouns, adjectives and adverbs e.g. *to grant a scholarship* while grammatical collocations usually concern one closed class word such as prepositions and

auxiliary words (Kimmes, 2004) e.g. *to turn on the light*.

Similarly, Herbst (1996) has also cited that Cowie (1994) and Robins (1971) have defined collocation as “the habitual association of a word in a language with other particular words in sentences such as *white coffee, green with jealousy* and *maiden speech*” (p. 383). Meanwhile, Nattinger and DeCarrico (1992) have elaborated that collocations are “strings of specific lexical items that co-occur with a mutual expectancy greater than chance (p. 36).” Woolard (2000) supports this definition citing that a collocation means “words which are statistically much more likely to appear together than random chance suggests” (p.29).

Also for Sinclair (1991), collocations are “ items that occur physically together or have stronger chances of being mentioned together” (p. 170). Normally, there is no specific rule or regulation to explain these co-occurrences. Sinclair has given examples of collocations as follows:

“We say to *break rules* but not to *break regulations*; to *hold a funeral* but not to *hold a burial*; to *make an attempt* but not to *have an attempt* and to *have a try* but not to *make a try*, etc” (p.170).

### Characteristics of collocations

The characteristics of a collocation can possibly be summarized as, firstly, a frequent

co-occurrence of items between which no word can be inserted. For example, in *bread and butter*, it is would be very unusual to add a word to this collocation, like it is unlikely to say *bread, cheese and butter*. It will definitely sound foreign if anyone happens to say so. Secondly, the components of a collocation cannot be replaced by a synonym or word of similar meaning (Yang & Hendricks, 2004).

For example (p.54):

*John is away on a business trip.*

*John is away on a business journey.*

It is clearly seen that *trip* is synonymous with *journey*; however, when using *journey* with *business*, it is unacceptable because *journey* does not collocate with *business*.

Thirdly, collocations are irreversible binomials (James, 1998). The order of the parts of a collocation is more or less fixed, for example, *bed and breakfast*, not *breakfast and bed*, or *slowly but surely, more or less, fish'n chips, for or against*, etc. (p. 72). However, it is not ungrammatical to put the parts the other way around. Finally, collocations have a high degree of predictability (Hill, 2000); for example, if you hear a collocation “*more or... ,*” you automatically expect that “*less*” will follow.

### Classifications of collocations

It seems that linguists and researchers have their own ways of classifying collocations. This

paper sets down how some of them have grouped collocations.

Hill (2000, pp. 63–64) has classified collocations as follows:

1. Unique collocations: These refer to collocations that are fixed and cannot be replaced by an/other word/s and are “highly predictable” (p. 51). Hill has given an example, *We shrug our shoulders* explaining that *shrug* is used only with *shoulders*, not with our *head* or any other parts of the body.

2. Strong collocations: These are not unique collocations but are strong or very strong. Strong collocations have a few other possible collocates. For example, the word *rancid* can be used with only a few words such as *butter* and *oil*.

3. Weak collocations: These refer to words that may have a number of word partners and can be easily predicted such as *dark green*, *light green*, *pale green*, *bright green*, *emerald green*, *lime green*, *lush green*, *rich green*, *olive green*, *dull green*, etc.

4. Medium–strength collocations: These refer to collocations that can sometimes be weak collocations; however, they are not common for EFL/ESL learners, such as *a door key* and *a key person*. Normally learners already know each individual word but not as the whole phrase. Hill has suggested that this kind of collocation should strongly be emphasized in class.

For James (1998), collocations are divided into three degrees as illustrated below:

1. Collocations of semantically determined word selection

This means the meaning of an item determines the appropriate word partnership; for example, it is right to say *state-of-the-art technology* but not *state-of-the-art manager* since a manager cannot be *state-of-the-art*.

2. Collocations of statistically weighted preferences

This degree of collocation is due to the frequency of the co-occurrence of words, or the preferences of the users of the language; for example, we can say he works *night and day* but he works *day and night* is preferred. Also *black and white*, and *knife and fork* are acceptable, but not *white and black*; or *fork and knife*.

3. Collocations combined arbitrarily

This kind of collocation conveys that there is no reason for the combination of words to form a collocation; for example, we say *to bake a cake*, but *to make pancakes*, not *to bake pancakes* or *to make an attempt* but *to have a try*, etc (p.152).

### **Factors causing EFL/ESL learners' collocational violations**

Various factors are said to contribute to EFL/ESL learners' collocational violations; however, it seems that the most interesting causes of errors are due to the learners' close

adherence to the source language and their application of the knowledge of the native language to the foreign language (Nesselhaul, 2003; Huang, 2001; Farghal & Obiedat, 1995; Baker, 1992 ; Zughoul, 1991; Corder, 1983).

Huang (2001), for example, has pinpointed that two major factors influence the performance in collocation for EFL learners. The first is their native language interference and this interference is greater when the learners do translation work. The other factor is the EFL learners' collocational competence. Huang has indicated that "Compared with their native-speaker counterparts, the ESL/EFL learners produced a lower percentage of conventional collocations but a higher percentage of deviant combinations" (Huang, 2001).

For Baker (1992), the sources of errors and problems are due to the effect of the collocation pattern of the source text, the misinterpretation of the source-language collocations, the decision to choose between accuracy and naturalness in translating collocations, culturebound factors, and marked collocations in the source text.

Meanwhile, Zughoul, (1991) has found that wrong collocations are caused by first, learners' inadequate experience in reading English resulting in low-level exposure to the English language, giving rise to particularly an inability in using correct collocations. Second, when translating from the source to the target language, learners rely heavily on bilingual dictionaries that only give a synonym without any

explanation or sample sentences. This can easily cause collocational violations; for example, a learner may write *a qualified hotel* in stead of *a quality hotel* after looking up the meaning in a bilingual dictionary. Third, learners' absorption in the source language can lead them to produce collocations that are "foreign" such as *a complimentary selling program* in stead of *a sales promotion*. Lastly, learners transfer the source language to the target language with a hypothesis that the two languages have similar collocations such as, in the Thai context, learners might produce *to open the light* in stead of *to turn on the light*.

To recap, most linguists and researchers agree that the first language greatly affects the production of collocations of EFL/ESL learners.

### **Raising awareness of English collocations among EFL/ESL personnel**

Making students aware of collocations is necessary. It is essential to make EFL/ESL learners realize the importance of collocations when learning new vocabulary, especially at the earliest stage of English learning. Hill (2000) has said, " Collocation should play an important part in our teaching from lesson one" (p.60). Teachers should start with easy simple collocations with young EFL/ESL learners and the level of collocation difficulty can be increased as time passes. Farghal and Obiedat (1995) have pointed out that "the highlighting of

collocation aspects of lexical items is as important as teaching them individually. This, we believe, is the sole way of nurturing the active use of language and helping the foreign learners construct lexically as well as grammatically acceptable sentences” ( p.3 ).

Wei (1999) has pointed out that in order that EFL/ESL learners reach the stage of correct production of collocations, they have to acquire an understanding of how words collocate with one another and they need to register a considerable variety of collocations. Most importantly, emphasis should be put on raising learners’ awareness of collocations because this could arouse them to search for collocations from books, dictionaries, and interactions with English native speakers.

## RESEARCH METHOD

The sample group consisted of 32 fourth-year English majors at Srinakharinwirot University in Bangkok. The data were taken from the 32 students’ translations of one out of three parts of the final examination on Business Translation. The selected part for the study was a business news article translated from Thai into English. The data, the Thai into English translated text, was parsed into 30 parts based on the Thai version. Each Thai parsed part was meaningful in itself.

For example,

*Source Language* : ราคาแพง

*Target Language* : high-priced or costly

*The student’s translation into the target language* : expensive price

Then, the parsed parts were listed and marked by two specialists who are native speakers of English. The specialists were given the same guidelines and explicitly instructed to be consistent in marking i.e. using the same criteria in marking the same data.

The study was restricted to the violated items relating to verbs, nouns, adjectives, adverbs and prepositions only (James 1998; Lewis 2000). Therefore, eight patterns of collocations were studied as follows: adjective + noun, verb + noun, noun + noun, verb + adverb/adverb + verb, adverb + adjective, noun + verb, verb + preposition, and preposition + noun. Violations on articles, spelling, punctuation, or other grammatical points were not considered in this study.

The raters’ reliability was studied and analyzed to assure the quality of the data as follows:

### 1.1. Item reliability

The percentage of congruency was studied to investigate the correlation in rating each item of the two raters. The result ranged from .53 to 1.00 or from 53 % to 100 % indicating that the congruency of the two raters’

marking was good and reliable.

The correlation of each translated parsed part was later investigated using the Phi correlation formula ( $\phi$ ) (Russell, 2000). The result ranged from .181 to 1.00 showing that the markings were well correlated. Nineteen parsed parts yielded at the significance levels of .01 and .05.

### 1.2. Test reliability

The correlation between the two scores of the same test by two independent raters was studied using Pearson Product Moment Correlation formula (Hopkins and Antes, 1990). The correlation was at 0.487, at the level of significance of .01.

### 1.3. Item analysis procedures

The item easiness index (Hopkins and Antes, 1990) was investigated to find out how the students did in the test. The index range was between .00 and .91.

The item discrimination index was investigated by means of item–total correlation using Pearson Product Moment Correlation formula (Hopkins and Antes, 1990). The index range was between .28 and .35.

Finally, the informants' collocational violations were analyzed and plausible explanations for the errors were made.

## FINDINGS

### 1. The students' ability in translating a business text from Thai into English

The students' English proficiency level was rather low since the mean was at 11.5 while the total score was 30. The standard deviation, however, was at 2.828 showing that the whole group's proficiency level was close to each other. The low level of the mean of the test score is possibly due to the rather difficult test. There were ten out of 30 parsed parts that the informants had found very difficult (the item–easiness index range: .00–.19), while three parsed parts were rather difficult (.20–.39).

The item–easiness index of the whole test ranged from .00 to .91. There were three out of 30 parsed parts that were very easy (.80–1.00). Only one item (.60 – .79) was rather easy.

Meanwhile, the item discrimination index of this examination was investigated by using Pearson Product Moment Correlation. The correlation ranged from .28 to .35 indicating that the examination was considered a good one and had a high discriminatory level.

### 2. Analysis of collocational violations

The found violations were classified into eight patterns with restriction to verbs, nouns, adjectives, adverbs and prepositions only (James 1998; Lewis 2000), as earlier mentioned. The investigation showed that the adjective + noun pattern of collocational violations were

found at the highest level (21.31%), followed by the patterns of verb + noun (18.03%), noun + verb (14.75%), noun + noun (13.11%), verb + adverb / adverb + verb (11.47%), and verb + preposition (11.47%) respectively. The pattern of preposition + noun collocations was found being equally violated at the lowest level (4.91%). However, violations of adverb + adjective collocations were not observed because students rarely used this pattern and, when they did, used it correctly.

### 3. Plausible explanations to account for the collocational violations

The eight patterns of collocational violations were analyzed. An attempt was made to give plausible explanations for the sources of violations by applying strategies and findings of various interesting linguists and researchers. The possible sources of violations are illustrated in the following table:

Possible Source of Violations	No. of Source Cited	Percentage
The informants' indulgence in over-literal translation (Corder, 1983) or the engrossing effect of the source text patterning (Baker, 1992)	19	32.76
The informants' application of the strategy of paraphrasing (Farghal & Obiedat, 1995; Huang, 2001)	6	10.34
The informants' low knowledge of grammatical collocations (Lewis, 2000)	6	10.34
The misinterpretation of the meaning of the source language collocations (Baker, 1992)	5	8.62
The informants' application of the strategy of synonymy (Farghal and Obiedat, 1995)	5	8.62
Learners' creative invention and the strategy of analogies (Granger, 1998 and Howarth, 1998, cited in Huang, 2001)	4	6.90
The informants' low collocational specialization (Nation, 2001)	4	6.90



Possible Source of Violations	No. of Source Cited	Percentage
The informants' application of the strategy of transferring (Farghal & Obiedat, 1995; Huang, 2001)	2	3.44
The informants' application of the strategy of avoidance (Farghal & Obiedat, 1995; Huang, 2001)	2	3.44
The informants' knowledge of restricted collocations (Howarth, 1998)	2	3.44
The informants' translation of marked collocations (Baker, 1992)	2	3.44
The informants' tension in translating between accuracy and naturalness(Baker, 1992)	1	1.72
Total	58	100

It is seen from the above table that the collocational violations studied in this specific research were possibly due to the informants' indulgence in over-literal translation and the engrossing effect of the source text patterning at the highest level (32.76%). It is also interesting to observe that the other sources of violations cited as plausible accounts for the errors were at very low levels (10.34%, 8.62%, 6.90%, 3.44% respectively). Finally, the informants' tension in translating between accuracy and naturalness was found to be at the lowest level (1.72%).

**DISCUSSION**

The results of the study reveal that the informants' ability in translating a business text from Thai into English, with effective use of

collocations, was at a low level. Deficiency of collocational knowledge resulted in erroneous productions of both lexical and grammatical collocations. The data obtained showed that the informants were not aware of collocations when translating a Thai text into English at all. In fact, the students were allowed to use all kinds of dictionaries, including a collocation dictionary, even in their examination. Although the use of this kind of dictionary had been encouraged in class, the plausible explanation was that not enough intensive practice was adopted.

The analysis of the violations of collocations showed that major causes of the violations were the learners' interference of the patterns of the

source language and their unawareness of collocations due to deficiency of collocational knowledge. It is plausible that the informants had a big gap between the receptive and productive knowledge of collocations (Farghal and Obiedat, 1995, cited in Huang). The students might have learned about new vocabulary with word partnerships but not in a meaningful way. Therefore, they could not produce correct collocations afterwards.

The results of this research support those of the previous studies conducted by several researchers, one of which was a study by Hsu (2002) regarding factors affecting the learnability of lexical collocations. He indicated that one of the factors was L1/L2 difference and it played a key role in producing effective collocations.

The findings of this research also endorse Huang (2001)'s study on Taiwanese EFL learners' knowledge of English collocations. Interestingly, the plausible explanations for the collocational violations committed by Thai learners found in this study revealed applications of a number of strategies which were similar to those of Huang's. The mentioned strategies were, for example, the strategies of transferring, avoidance, paraphrasing, and synonymy. The overall result indicated that EFL learners possessed insufficient collocational knowledge. Huang demonstrated that the problems were probably due to "the habit of learning English vocabulary as isolated words" and unawareness of word partnerships. When teaching vocabulary to ESL learners, it was recommended to integrate the teaching of collocation by including "cultural data,

metaphorical meanings and the historical origins associated (Huang, 2001)" to reach an ultimate goal of effective communication.

Similarly, the findings of Farghal and Obiedat (1995) support this research. They found out that Arabic ESL learners and student translators had a problem of collocational deficiency and were unaware of collocations, resulting in their inability to communicate with effective collocations. Also, the data analysis revealed that in the translation process, the informants used the following strategies: synonymy, avoidance, transferring and paraphrasing. The two researchers recommended that the only way for ESL/EFL learners to be proficient in the English language was to "highlight the collocational aspects of lexical items." They added that collocations should be focused on seriously among L2 learners and teachers alike and that "foreign language syllabuses and language instructors should single collocations out as the most needed and useful genre of prefabricated speech" (Farghal and Obiedat, 1995).

This study has identified common areas of collocational violations and through the application of various specialists' strategies attempted to offer plausible explanations. EFL/ESL personnel's awareness of collocations has been raised, bearing in mind that "Lexical items are central to language use and should be central to language teaching (Zimmerman, 1997, p. 89) and that collocational competence is one of the major means for being successful in EFL/ESL pedagogy."

## REFERENCES

- Baker, M. (1992). *In other words*. London: Routledge.
- Bahn, J & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, 21, 101 – 114.
- Benson, M. (1985). Collocations and Idioms. In Ilson, R. (Ed.), *Dictionaries, lexicography and language learning*. Oxford: Pergamon Press.
- Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford: Oxford University Press.
- Cowie, A.P. (1994). Phraseology. In R.E. Asher (Ed.) *The encyclopaedia of language and linguistics*, 6 (3168–71). Oxford: Oxford University Press.
- Farghal, M. & Obiedat, H. (1995). Collocations: a neglected variable in EFL writings, *IRAL: International Review of Applied Linguistics in Language Teaching*, 33, 315–331. Retrieved September 2003 from ProQuest Direct (No.0019042X)
- Herbst, T. (1996). What are collocations: sandy beaches or false teeth? *English Studies*, 4, 379–393.
- Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation: Further development in the lexical approach* (pp.47–69). Oxford: Oxford University Press.
- Hopkins, C. D. & Antes, R. L. (1990). *Classroom measurement and evaluation*. (3<sup>rd</sup>ed.) Itasca: F.E. Peacock Publishers.
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19, 45–72.
- Hsu, J. (2002). *Development of collocational proficiency in a workshop on English for general business purposes for Taiwanese college students, China* (Abstract). Pennsylvania: Indiana University of Pennsylvania. Retrieved August 22, 2003 from Pro Quest (No. 3040680)
- Huang, L.(2001). *Knowledge of English collocations: an analysis of Taiwanese EFL learners*. Retrieved July 22, 2003 from the World Wide Web: [www.utexas.edu/students/flesa/tpfle/contents7.doc](http://www.utexas.edu/students/flesa/tpfle/contents7.doc)
- James, C. (1998). *Errors in language learning and use: Exploring error analysis*. New York: Addison Wesley Longman Limited.

- Kimmes, A.M. *An investigation of the usage and collocability of English verbs of thinking based on the online edition of The New York Time* (Abstract). New Haven: Southern Connecticut State University. Retrieved April 3, 2005 From Pro Quest (No. 1420044)
- Kumiko, O., Kengo, S. & Masakazu, N. (2001). Extracting translation uninterrupted collocations from bilingual corpora. (Abstract No. 122-003). *IPSJ SIG Notes Natural Language*. Retrieved August 21, 2003 from the World Wide Web: <http://www.ipsj.or.jp/members/SIGNotes/Eng/01/1997/122/article003.html>.
- Lewis, M. (2000). Language in the lexical approach. In M. Lewis (Ed.), *Teaching collocation: Further development in the lexical approach* (pp.126-154). Oxford: Oxford University Press.
- Lewis, M. (Ed.) (2000). *Teaching collocation: Further development in the lexical approach*. Oxford: Oxford University Press.
- Nation, I.S.P.(2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nattinger, J. R. and DeCarrico, J.S. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 223. Retrieved September 2003 from ProQuest Direct (ISSN No.1426001)
- Park, C-J, Lee, J-H, Lee, G & Kakehi, K. (2001). Collocation-based transfer method in Japanese-Korean machine translation. (Abstract). *IPSJ Journal*. Retrieved July 22, 2003 from the World Wide Web: <http://www.ipsj.or.jp/members/Journal/Eng/3804/article005.html>.
- Russel, B. H. (2000). *Social research methods: qualitative and quantitative approaches*. California: Sage Publications Inc.
- Sinclair, J. (1991). *Corpus concordance collocation*. Hong Kong: Oxford University Press.
- Wei, Y. (1999, March). *Teaching collocations for productive vocabulary development*. Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages. New York, NY.
- Woolard, G. (2000). Collocation – encouraging learner independence. In M. Lewis (Ed.), *Teaching collocation: Further development in the lexical approach* (pp.28- 46). Oxford: Oxford University Press.

- Yang, Y. & Hendricks, A. (2004). Collocation awareness in the writing process. *Reflections of English Language Teaching*, 3, 51-78
- Yu-Chih, S & Li-Yuch, W. (2003). Concordancers in the EFL classroom: cognitive approaches and collocation difficulty. *Computer Assisted Learning*, 16, 83-94.
- Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition*. Cambridge: Cambridge University Press.
- Zughoul, M. R. (1991). Lexical choice: towards writing problematic word lists. *IRAL: International Review of Applied Linguistics in Language Teaching*, Retrieved August 26, 2003 from EBSCO host Research Databases. (No.0019042X)