

Collocational Errors in EFL Learners' Interlanguage

Supakorn Phoocharoensil

Language Institute, Thammasat University
English Language Center Building
2 Phrachan Rd., Phrabarommaharajchawang, Phranakorn, Bangkok
Thailand 10200
Tel: +66 (0)2613-3101 - 3 E-mail: yhee143@gmail.com

The research is financed by a scholarship provided by the Language Institute of Thammasat University.

Abstract

This research study is aimed at examining Thai EFL learners' errors in their acquisition of English collocations. The data, drawn from essays written by two groups of participants differing in L2 proficiency, indicate actual problems with which the learners are really faced. As regards the origins of these collocational errors, first language transfer seems to be the most outstanding strategy which they adopted. Where the collocations in Thai and English appear incongruent, deviations often arise. The L1-based errors, evidenced by the data, deal with preposition addition, preposition omission, incorrect word choice, and collocate redundancy. It comes as a surprise to discover the high-proficiency learners' heavy reliance on the native language, to which those with low proficiency are expected to resort. In addition to L1 transfer, the learners also apply synonymy and overgeneralization, both of which contribute to collocational deviations in the target language.

Keywords: Collocation, Interlanguage, Second Language Acquisition, Error, EFL learner, Learning strategy, Learners' problem

1. Introduction

It has been discovered that learners of English as a foreign language (EFL), despite having a large number of words in the second language (L2), are not fully capable of putting words together in a way that native speakers naturally do. In other words, EFL learners can be said to lack collocational competence in L2. Collocation, i.e. how words tend to co-occur in a language, seems to cause quite a few problems for English learners since even though two words, when placed together, may appear to seemingly make sense, the resulting combination is sometimes regarded as a deviation in the target language. This probably implies that collocation is arbitrary for language learners (Hill, 2000).

Several past studies investigating second language acquisition of English collocations have shown that EFL learners' problems are due to different factors (e.g. Fan, 2009; Huang, 2001). One of the major reasons for collocational errors pertains to native language influence (e.g. Bahns, 1993; Bahns & Eldaw, 1993; Fan, 2009; Koya, 2003; Nesselhauf, 2003, 2005; Ying, 2009). Some have also found that learners rely on certain learning strategies, such as synonymy (e.g. Boonyasaquan, 2006; Farghal & Obiedat, 1995; Mongkolchai, 2008), repetition and overgeneralization (e.g. Fan, 2009; Granger, 1998; Howarth, 1998; Shih, 2000), etc.

However, most of the previous research studies mentioned above focused on one single group of EFL learners, particularly those with high proficiency (e.g. Boonyasaquan, 2006; Mongkolchai, 2008). Hardly do such studies reflect a true picture of how learners' collocational knowledge has been developed through time. To put it another way, there has been little research to date examining learners' errors in the total interlanguage system with respect to collocation learning. It is for this reason that the present study was undertaken to bridge the gap, i.e. to scrutinize the collocational competence of high-proficiency and low-proficiency Thai learners of English.

The definitions of collocation, the concepts of interlanguage, including Contrastive Analysis and Error Analysis, and related past studies are reviewed in the next section.

2. Literature review

2.1 Collocation

2.1.1 Defining collocation

The term collocation has been defined by different scholars but in a similar fashion. It was introduced by Firth (1957) to refer to a combination of words associated with each other. Sinclair (1991) defines collocations as “items that occur physically together or have strong chances of being mentioned together” (p. 170). According to Lewis (1997), collocation is defined as “the readily observable phenomenon whereby certain words co-occur in natural text with greater than random frequency” (p. 8). The occurrence of collocation is statistically significant (Lewis, 2000). In support of Lewis (1997, 2000), Hill (2000) suggests that “collocation is a predictable combination of words” (p. 51). Examples of common English collocations are *rancid butter*, *make a decision*, *Internet access* etc. Hill (2000) also noticed that some collocations are fixed and highly predictable from one of the component words. For instance, the verb *shrug* apparently almost always co-selects the noun *one’s shoulder* as its neighboring word, i.e. a collocate. In this way, *to shrug one’s shoulder* can be viewed as a strong or restricted collocation. In contrast, some collocations are considered so weak that their occurrences often go unnoticed or seem too general as the two component words are inclined to occur freely. For example, the adjective *good* can co-occur with a tremendous variety of noun collocates, e.g. *a good boy*, *a good school*, *a good teacher*, *good food*, etc., all of which are considered weak collocations.

2.1.2 Classification of collocations

A number of linguists have proposed certain criteria so far for distinguishing different kinds of collocations. The current study has adopted Benson, Benson, & Ilson (1997)’s collocation classification: lexical collocations and grammatical collocations.

2.1.2.1 Lexical collocations

Lexical collocations are composed of two or more content words, i.e. nouns, verbs, adjectives, and adverbs. Examples of this kind of collocation is presented below:

adjective + noun	:	sour milk
verb + noun	:	conduct research
noun + verb	:	dust accumulates
adverb + adjective	:	mentally disabled
verb + adverb or	:	move freely
adverb + verb	:	proudly present

2.1.2.2 Grammatical collocations

Grammatical collocations refer to combinations comprising a content word and a function word, which is usually a preposition, as illustrated below:

noun + preposition	:	an increase in
verb + preposition	:	elaborate on
adjective + preposition	:	familiar with
preposition + noun	:	on probation

2.1.3 Collocation learning strategies and Relevant L2 studies

There appear many different strategies applied by language learners in their attempt to acquire L2 collocations. This paper summarizes only some major learning strategies that often lead learners to collocational errors in L2 English: first language transfer, synonymy, and repetition and overgeneralization.

2.1.3.1 First language transfer

Learners' native language (L1) largely has an impact on their subsequent learning of L2 collocations (Nesselhauf, 2003, 2005). Learners' reliance upon their L1 collocational knowledge may represent their assumption that there is a one-to-one correspondence between L1 and L2 collocational choices. Fortunately, where there is an exactly identical match between collocations in both languages, transfer from learners' mother tongue could result in positive, satisfactory production (Ellis, 2008). For instance, the combination *in reality* appears to be possible in both Thai and English. As a result, it is very likely that Thai learners will become successful in transferring this particular collocation from L1 Thai to L2 English.

Nonetheless, such success based on native language transfer is not always the case (Nesselhauf, 2003, 2005). Discrepancies between L1 and L2 collocations can also cause some problems for EFL learners. That is, whenever collocations in the mother tongue and the target language do not match, deviant collocational structures often arise. According to several previous studies, native language influence is noticeable in EFL learners' collocations. By and large, there obviously exists negative transfer from L1. Bisk-up (1992), in investigating Polish and German EFL learners' performance in English collocation use, revealed that the learners, based on risk taking, did transfer their L1 collocational knowledge to their production of L2 collocations, thus evidently leading to erroneous use of English collocations. For example, while the targetlike collocation in English is *to set a record*, the Polish learners tended to use *to state a record*, which is indicative of an L1 collocational pattern. Likewise, the German learners were found to produce the L1-based deviation *to lend a bookshop* instead of the targetlike version *to run a bookshop*.

In a similar way, Bahns (1993) and Bahns & Eldaw (1993) reported on the role of mother tongue in English collocation acquisition. That is, German learners of English, in a translation task from German to English, were found to be successful with transferring from L1 collocational knowledge when L2 collocations have L1 equivalents. However, negative transfer was also remarkable when there appears non-congruence between collocations in both languages. L1 interference can be seen in Huang (2001) as well when Taiwanese EFL university students, having been asked to do a sentence-completion test, created L2 combinations based on L1, such as *a black horse* rather than the target-like collocation *a dark horse*.

Nesselhauf (2003) provided support for the previously mentioned studies in that L1 influence, in her study of collocations used by German EFL learners, is considerable, resulting in L2 errors for several times. She also confirmed the significance of native language impact on L2 collocation learning, suggesting that since L1-L2 collocational incompatibility is a major source of errors in learner language, English teachers should concentrate on such non-congruent collocations in the two languages in order to prevent learners from committing such transfer errors.

It is also worth noticing that in Koya (2003), even high-proficiency students seem to heavily rely on their knowledge of L1 collocations, which came as a surprise to the researcher himself since he had predicted to see far less evidence of L1 transfer in this group of high-proficiency students. On the other hand, low-proficiency learners were found to apply an avoidance strategy and astoundingly depended less on their first language. This supplies counter evidence against much past literature which indicated that L1 transfer is characteristic of low-proficiency learners (e.g. Ellis, 1987; Odlin, 1989).

In addition, Fan (2009), in an examination of Hong Kong ESL learners' collocation production in writing, also discovered an adverse effect that L1 Chinese had on the participants' use of English collocations. In particular, the study found non-standard L2 collocations that seem to result from word-for-word translation from Chinese, such as *as left/right face* or *left side face*, which are not present in native speakers' corpora (Fan, 2009, p. 118).

Another study that is in line with the aforementioned ones as to L1 transfer is Ying (2009). In the study of English collocations produced by Chinese speakers, i.e. English majors and non-English majors, Ying found that collocations which have no translation equivalents in L1 are considered difficult, in comparison to those which are congruent with L1. In more details, the learners probably searched for L1 equivalents with no awareness of L1-L2 incongruity and then produced L2 deviant combinations, which accords with Nesselhauf (2003). Moreover, for both groups of learners, errors in lexical collocations clearly outnumber

those in grammatical collocations.

With respect to research on L2 acquisition of English collocations by Thai EFL learners, L1 transfer has also been prevalent. As revealed by Boonyasaquan (2006), who investigated Thai learners' collocational violations in translating a Thai business article into English, native language, i.e. Thai, was one of the most common sources of errors. The learners seemed to directly translate some L1 collocations into non-targetlike English. An example given in this study is **expensive price*, which is assumed to be directly translated from *raakhaa pej* 'high price', i.e. literally *raakhaa* 'price' and *pej* 'expensive'. Another study the results of which are consistent with Boonyasaquan (2006) in terms of first language influence on collocation use is Mongkolchai (2008), who studied English collocational competence of third-year English majors in a university in Bangkok, Thailand. The research instruments, viz. a sentence completion test and a multiple-choice test, seemed to elicit collocational data presenting evidence of native language transfer. A clear example of transfer-based errors is the use of the preposition *in* in place of *at* in the sentence **I became skillful in drawing* (Mongkolchai, 2008, p. 46)

It is evident that interference from EFL learners' native language plays a crucial role in L2 collocation acquisition. As shown in the aforementioned studies, the first language seems to have a negative effect on their use of English collocations, resulting in L2 erroneous combinations. In 2.1.3.2, another learning strategy also causing problems for EFL students, i.e. synonymy, is discussed.

2.1.3.2 Synonymy strategy

Aside from dependence on their native tongue, EFL learners in the process of learning collocations are sometimes seen to adopt an analogy strategy referred to as synonymy strategy. This is often used by learners whose L2 proficiency is limited. They may try substituting a synonym for a word in L2, unaware of constituting a collocational violation. In actuality, a very limited number of synonyms in English can occur in the same grammatical pattern (Nation, 2001). In other words, words that are very close in meaning do not always share the same grammatical collocation. For instance, even though the verbs *ask* and *plead* are semantically similar, i.e. involving making a request (Cambridge advanced learners' dictionary, 2008, pp. 74-75 & 1085), the grammatical patterns in which the verbs are likely to occur are different. That is, the verb *ask* is used in the pattern *ask someone + infinitive with to*, whereas the verb *plead* requires the preposition *with*, as in *plead with someone + infinitive with to*. For this reason, a substitution of *plead* for *ask* in the grammatical pattern of the latter verb, i.e. without *with*, causes ungrammaticality in English (Phoocharoensil, 2010, p. 242).

According to many studies of L2 English collocation acquisition, synonymy has appeared to be a common learning strategy. In Farghal & Obiedat (1995), it was indicated that Arabic EFL learners greatly relied on the open-choice principle for word selection, replacing a word with its synonym. Such a strategy often led them to deviant, ungrammatical collocations in English. In a similar vein, Howarth (1996, 1998) demonstrated that L2 learners seemed to draw an analogy between collocates of two synonyms, thus frequently resulting in errors in the target language. For example, they produced the deviant combination **adopt ways*, which was presumably caused by analogy with the correct collocation *adopt an approach* (Howarth, 1998, p. 41).

Like the above studies having been mentioned, Zughol & Abdul-Fattah (2001) discovered assumed synonymy in the use of English collocations by Arabic speakers. It was reported that as a consequence of the nature of the instructional input the learners received in class and the impact of bilingual dictionaries, the learners' collocation use was evidently based on a synonymy strategy, which violates the selectional restrictions, i.e. semantic constraints, of the target language. For instance, the verb *failed* was incorrectly employed as opposed to *defeated* in the sentence **The enemy was failed in the battle* (Zughol & Abdul-Fattah, 2001, p. 11)

As regards some research studies on Thai learners' acquisition of English collocations, synonymy has also been discovered. Boonyasaquan (2006), in her study of how Thai EFL learners translated a business news articles from Thai to English, reported on their collocational deviations stemming from an application of synonymy, which accounts for 8.62% of all the collocational errors. A clear example given in the study is

*a *qualified hotel* in lieu of a *quality hotel*, which may reflect the learners' confusion over the use of the synonyms *qualified* and *quality* (Boonyasaquan, 2006, p. 83). Lending support to Boonyasaquan (2006), Mongkolchai (2008) noticed collocational errors committed by English majors, who speak L1 Thai, through synonymy. For instance, they used *a *newspaper booth*, which is a deviation from the targetlike one a *newspaper kiosk*, commonly found in a native speaker corpus.

2.1.3.3 Repetition and Overgeneralization

Repetition is another strategy on which EFL collocation learners, especially those with a low level of L2 collocational knowledge, often depend. With repetition, learners resort to the repeated use of a limited number of familiar collocations. Moreover, repetition may occur as a result of learners' lack of confidence to create L2 collocations through analogy (Howarth, 1998). Put simply, they may not want to risk using L2 combinations with which they are unfamiliar. Research studies on L2 English collocation learning have pinpointed repetition as an origin of collocational problems.

Among those studies, Granger (1998) shows that French learners of English tended to repeatedly employ the intensifier *very* in the combination of adverb + adjective. Furthermore, some other collocations, e.g. *deeply-rooted*, recursively occurred in their writing as well. Granger, Paquot & Rayson (2006) corroborated Granger (1998) in that EFL learners seem to overuse a limited group of collocations perhaps because they stick to familiar formulaic sequences which they feel safe to use. Similarly, Howarth (1998) also discovered a difference between individual writers' repetition of conventional collocations. Shih (2000) was devoted to an investigation of overused collocations in a Taiwanese learner corpus of English, focusing on a set of synonyms *big*, *large*, and *great*. The findings from a comparative study of Taiwanese Learner Corpus of English and British National Corpus (BNC) showed that the collocations with *big* were significantly overused by Taiwanese learners. More precisely, the learners used *big* far more frequently than native speakers normally do when describing abstract concepts, whereas the use of *big* referring to concrete objects occurs with more frequency in the native speaker corpus. Shih posited that repetition is viewed as a simplification strategy or overgeneralization applied by Taiwanese learners when faced with L2 collocational problems. In other words, the word *big* is perhaps extended to abstract concepts, which is not a normal practice of native speakers'.

According to Zughol & Abdul-Fattah (2001), overgeneralization, i.e. the extension of the use of a certain L2 feature to another, has been found as a source of incorrect use of L2 English collocations, and this strategy is viewed as a characteristic of learner language. The subjects in this study confused the words *shame* and *ashamed*, thereby extending the use of *ashamed*, while the word *shame* was intended.

In 2.2, the concept of interlanguage, together with Contrastive and Error Analyses, is reviewed.

2.2 Interlanguage

2.2.1 Contrastive Analysis and Error Analysis

2.2.1.1 Contrastive Analysis (CA)

To clearly understand the concept of interlanguage (IL), it is useful to look at how the term has been coined and developed over time. Before discussing the major characteristics of IL, the notions of Contrastive Analysis (CA) and Error Analysis (EA) should be taken into consideration. With respect to the Contrastive Analysis Hypothesis (CAH), Lado (1957) proposed that language learning is similar to habit formation. Learning a new language is equated with replacing an old, existing habit with a new one. The major goal of contrastive analysts, who compare and contrast linguistic features in the first language (L1) and the second language (L2), is to predict the area of difficulty with which L2 learners will be confronted in learning L2 or the target language (TL). It is assumed that where L1 and L2 are alike, rarely will difficulty arise; however, where there are differences between the two languages, a more tendency of learners' error production as well as reception should be seen. That is, learners' native language is viewed as the major source of error in L2 learning. Pedagogically speaking, learners are expected to learn the L1-L2 differences in order for them to avoid error production in L2 (Gass & Selinker, 2008).

There appeared criticisms, nonetheless, of the CAH that gathered force in the early 1970s. First, the error

predictability of CA was in doubt. Certain scholars began to discover that some errors predicted by CA did not occur in reality. By contrast, some types of actual errors had never been predicted. In other words, such errors are hypothetical (Ellis, 1985). It has been demonstrated that learners' errors are not solely attributed to the differences between the native language and the target language but can also be due to their creativity to formulate L2 rules without being aware of certain exceptions (Chomsky, 1959). In short, L2 learners in the process of TL development may overgeneralize TL rules, ending up with ill-formedness. The drawbacks of CA apparently account for why the predictive power of CA was tremendously challenged.

2.2.1.2 Error Analysis (EA)

The shortcomings of CA as outlined above gave rise to Error Analysis (EA). Focusing on the actual errors learners make, EA aims at a comparison between the errors occurring in producing the TL and the TL form, in contrast to CA, where the comparison is made with the mother tongue (Gass & Selinker, 2008).

In his seminal article titled *The Significance of Learners' Errors*, Corder (1967) clearly distinguished between mistakes and errors. Mistakes are temporary and not persistent, made due to some performance factors, such as sleepiness, exhaustion, memory lapses, etc. When pointed out, mistakes can be corrected by learners themselves. On the other hand, errors are likely to occur repeatedly; they are not recognized by learners as errors. Put differently, errors are systematic. Hence, it is an error rather than a mistake that needs to be rectified for learners.

EA researchers generally follow more or less the same steps. To begin with, they collect language data containing written or oral errors and then identify these errors facing learners. After this, the errors are classified into different types before they are quantified. The next important step is to explain the possible cause(s) of the errors. Moreover, if the purpose of the EA is pedagogic, errors should also be evaluated, which means the seriousness of each error type has to be assessed so that L2 curriculum developers will be able to know which types of error should be highlighted in the teaching materials (Ellis, 1985, 2008).

Even though EA proves useful in language pedagogy, some limitations have been noticed. Bley-Vroman (1983) remarked that EA is guilty of comparative fallacy since it seeks to compare learner language and the target language. That is, it accounts for learner language solely in terms of target-language norms. In reality, EA proponents ignore the fact that learners create their own unique rule systems in the process of learning an L2. Such entire systems, commonly known as interlanguage, rather than merely errors are worth being analyzed (Selinker, 1992).

In connection with the notion of interlanguage, another drawback of EA is that it "fails to provide a complete picture of learner language." (Ellis, 2008, p. 61). Instead of investigating only the produced errors in L2 learning, it is advisable that the whole linguistic system created by learners be examined. In a nutshell, the significance of scrutinizing the totality of the learner's production should be recognized.

Additionally, most of EA studies look at learner language at a single point in time. Little care, in several cases, has been taken to separate out the errors which learners produce at different stages of development. For this reason, EA has not proved very effective in enabling us to understand learners' development of L2 knowledge over time. To put it another way, as Ellis (2008) suggested, an examination of how learners' errors change from one stage to another can shed light on the process of L2 acquisition.

Some of the problems of EA being shown, language researchers have been more and more interested in how learners develop linguistic competence in L2 as their own unique system, i.e. interlanguage, which will be discussed in 2.2.1.3.

2.2.2 Interlanguage Analysis

To fully comprehend the linguistic system of learner language rather than learners' imperfect language use, i.e. errors, interlanguage analysis is preferable (Ellis, 2008). The term interlanguage (IL) (Note 1), coined by Selinker (1972), refers to the systematic knowledge of an L2 which is independent of both learners' native language and the target language, but is linked to both L1 and L2 by interlingual identifications in the perception of learners (Tarone, 2006).

Central to the concept of IL is the notion of fossilization, defined as the cessation of L2 learning. Many L2

learners are found not to be capable of reaching TL competence. That is, they do not seem to reach the end of the IL continuum, stopping learning TL when their IL contains some rules different from those of the TL system (Selinker, 1972, 1992). According to McLaughlin (1987, p. 63), learners in the state of fossilization cease to “elaborate the interlanguage in some respect, no matter how long there is exposure, new data, or new teaching”. Fossilized structures are often realized as errors when learners have reached a stage in which a target feature still does not have the same form as the target language (Ellis, 1985).

In addition, the term interlanguage, according to Ellis (1985, 2008), reflects two different but related concepts. First, IL refers to the series of interlocking systems which characterize acquisition. In this way, learner language is regarded as the interlanguage continuum. Second, IL refers to “the structured system which the learner constructs at any given stage in his development” (Ellis, 1985, p. 47), i.e. an interlanguage.

Selinker (1972, 1992) stated five central cognitive processes responsible for shaping IL:

- a. *Native language transfer*: Learners’ L1 has some influence on the IL development. Such transfer can either facilitate or impede L2 learning.
- b. *Overgeneralization of TL rules*: Learners may extend a general L2 rule to the extent not covered by that rule. In other words, they have mastered a general rule but do not yet know all exceptions to the rule.
- c. *Transfer of training*: Learners apply rule which they have learned from teachers or textbooks. This rule application is sometimes successful when the resulting IL rule is identical to the TL rule. However, for some other times, erroneous use could result from previous training provided that the textbook or the instruction contains incorrect information.
- d. *Strategies of L2 learning*: Strategies are consciously applied in a bid to master the TL. For example, learners may use mnemonics to help them memorize target vocabulary. While learning strategies are frequently effective, they can sometimes lead to errors if learners are confused over their own created strategies.
- e. *Strategies of L2 communication*: When learners find that the linguistic item necessary for their communication in TL is unavailable, they may resort to various strategies of communication in order to get the meaning across. As a result, the linguistic forms produced in such attempts may become permanent in the learners’ IL system.

These above cognitive processes account for how learners’ IL develops. Accordingly, an investigation of these processes enables IL researchers to see a clearer overall picture of learner language. The following section discusses relevant previous studies on EFL learners’ acquisition of English collocations.

3. Research questions and Hypotheses

The present research study is aimed at providing answers for the following questions:

1. What are the collocational errors found in Thai EFL learners’ interlanguage?
2. What are the learning strategies on which Thai EFL learners depend in their English collocation learning?

In response to the above research questions, the hypotheses below were therefore formulated:

Hypothesis 1:

Thai EFL learners commit significantly more errors on lexical collocations than grammatical counterparts.

Hypothesis 2:

Thai EFL learners confronted with L2 English collocations rely on some learning strategies, such as native language transfer, synonymy, and overgeneralization.

4. Research methodology

4.1 Data collection

4.1.1 Participants

The participants in this study were 90 first-year undergraduate Thai EFL students at a university in Thailand, divided into two equal groups, i.e. high (H) and low (L), according to English proficiency. The students from various faculties were assigned to different classes based on their English O-NET scores, i.e. the University Entrance Exam under the Ministry of Education of Thailand. Regarding the cut-off scores, the participants whose O-NET scores range from 69-80 were assigned to Foundation English Course 3, i.e. the high group, whereas those whose scores are between 32 and 56 registered for English Foundation Course 2, i.e. the low group (Note 2). The score ranges above have been determined by the Registration Office of the university. The participants' native language is Thai. They have learned English as a foreign language (EFL) for at least 12 years, i.e. from grades 1-12.

4.1.2 Elicitation task

So as to elicit authentic interlanguage data from the participants, essay writing was chosen as the research tool. The two groups of participants were asked to write a descriptive essay, which is expected to effectively draw various types of collocation since this particular written task contains a wealth of modifications usually formed by word combinations (Langan, 2009), in class within 60 minutes. The researcher also acted as the teacher to facilitate the elicitation procedure and it was also hoped that the participants would be very cooperative and willing to do this task because they should be more familiar with the teacher than a language researcher in general. The participants were also informed that this written task would never be assessed according to grammatical accuracy. Consequently, they should feel relaxed enough to naturally produce the linguistic data that best represent their genuine collocational competence.

4.2 Data analysis

The data were then analyzed both quantitatively and qualitatively. The framework used in the current study was based on Benson et al. (1987), who have proposed two main types of collocations: lexical and grammatical collocations. To be more precise, four types of lexical collocations were the focus of this research project, i.e. noun + verb, verb + noun, adjective + noun, and verb + adverb (or adverb + verb). As for grammatical collocations, four types were analyzed: noun + preposition, verb + preposition, adjective + preposition, and preposition + noun. Only the incorrect collocations were examined. Furthermore, a combination would be labeled as deviant or erroneous provided that it does not conform to at least one of the following: 1.) four collocation dictionaries, i.e. The BBI Dictionary of English Word Combinations (1997), The LTP Dictionary of Selected Collocations (1997), Oxford Collocations Dictionary (2009), and Macmillan Collocations Dictionary (2010); 2.) the British National Corpus (BNC), which is a 100 million word collection of samples of written and spoken language from a wide range of sources, designed to represent a wide cross-section of British English from the later part of the 20th century (<http://www.natcorp.ox.ac.uk/>); and 3.) native speakers of British English and American English, both of whom are EFL teachers.

The entire number of deviant collocations produced by the two proficiency groups was counted and then compared. Qualitatively speaking, each type of collocational error was explained in relation to the learning strategies which the participants seemed to apply.

5. Findings and Discussion

The data from the high-proficiency participants, as shown in Table 1, reveal that the lexical collocational errors slightly outnumber the grammatical ones. With respect to the deviations in the lexical collocations, the high-proficiency learners appeared to be troubled most with verb + noun combinations (25.28%); they produced errors in adjective + noun collocations with the second highest frequency (13.79%). The problem caused by verb + adverb (or adverb + verb) combinations is ranked third (6.90%), while noun + verb combinations seemed to be the least problematic (5.75%).

As regards the grammatical collocations, the learners with high proficiency produced errors in four combinations. To be specific, they equally had problems with verb + preposition and preposition + noun

collocations (18.39%) with the highest frequency. Second to these two types in frequency is adjective + preposition combinations (8.05%). Additionally, noun + preposition collocations occurred with the lowest frequency (3.45%).

According to the data drawn from the low-proficiency learners' writing, demonstrated in Table 2, the errors in lexical collocations (58.56%) are obviously higher than those in grammatical collocations (41.44%). Regarding the deviant lexical collocations, the low-proficiency learners were found to have the most problems with verb + noun collocations (40.54%), followed by errors in noun + verb collocations (11.71%) and the adjective + noun collocations (5.41%) respectively. Verb + adverb collocation seemed to be the least troublesome (0.90%).

With regard to grammatical collocational errors, the most problematic type was the verb + preposition collocations (23.42%). The second most frequent type was the preposition + noun (12.61%), followed by the adjective + preposition (5.41%). No error in the noun+ preposition collocation was noticed.

All in all, as can be seen in the high-proficiency learners' data, the collocational errors in lexical types (51.72%) outnumber those in grammatical ones (48.28%). In the same manner, the data from the low-proficiency learners indicate the preponderance of lexical collocational deviations (58.56%) over the grammatical counterparts (41.44 %). These results support the first research hypothesis, which claims that there exist more problems of lexical collocations than those of grammatical ones. This also accords with past studies (e.g. Ying, 2009)

In 5.1, some examples of miscollocations found in the learner corpus, as well as possible sources of such collocational misuse, are given.

5.1 Sources of collocational errors

According to the linguistic data drawn from the two groups of students differing in their proficiency levels, it was very likely that certain learning strategies associated with cognitive processes (Selinker, 1992) were applied in an attempt to use English collocations, most of which apparently lead to erroneous combinations in the target language. The most prominent strategy upon which they seemed to depend was their native language, i.e. Thai. Additionally, some appeared to use synonymy and overgeneralization, to be discussed in detail below.

5.1.1 Native language transfer

Thai learners of English evidently heavily relied on their L1 collocational knowledge, transferring an equivalent from Thai to English. Unfortunately, where there are collocational mismatches between Thai and English, errors possibly arise. The first problem that may be related to L1 influence concerns preposition use. The learners were found to omit an obligatory English preposition after a verb probably because they relied on literal translation from the corresponding verb in Thai which is transitive, as in (1)-(2).

(1) * I *listen music* all the day without feeling bored in my room.

(2) * He always *takes care me* when I need sleep.

As for (1), the error occurs because the verb *listen*, always requiring the prepositional phrase introduced by *to*, cannot be immediately followed by a noun, e.g. *music*. By contrast, the verb *faj* 'listen' in Thai subcategorizes for a direct object. Thai learners may transfer this L1 knowledge to their L2 collocation production. In (2), the deviation stems from a deficiency of the preposition *of* as part of the verb phrase *take care of me*. As a matter of fact, the verb *duulee* 'take care of' in Thai is a transitive verb, so no following preposition is needed. This may cause the learners to omit the preposition *to* in the target English.

On the other hand, the participants in this study also seemed to unnecessarily add an extra preposition to a verb. This problem occurs when a verb in Thai requires a prepositional phrase to follow, whereas its English counterpart subcategorizes for a direct object. Such a discrepancy may account for the learners' preposition insertion in L2 English, as in (3) and (4).

(3) * I never *left from* my house.

(4) * I was to *leave from* home in short time.

The miscollocations in both (3) and (4) emanate from an existence of the preposition *from* next to the verb *left* and *leave* in (3) and (4) respectively. In English, *leave* is a transitive verb, so no object should follow this verb. The erroneous combination of **leave from* may originate from the Thai equivalent *จ้ะง* ‘leave’, which always needs the preposition *ค้ะง* ‘from’.

Another type of L1-transfer collocational problems connected with preposition use is an incorrect choice of preposition, as shown in (5)-(7) below.

(5) * When you’re *tired with* working and studying, you may need to take a rest.

(6) * I am really *impressed in* that special time.

(7) * I am *close with* paddy field and farmers’ lives.

In (5), the adjective + preposition collocation **tired with* is rather uncommon since the correct, widely-used preposition after *tired* should be *of*. This error is probably influenced by the Thai equivalent *น้ะง* ‘tired’, which is normally followed by the preposition *ค้ะง* ‘with’. Similarly, the problematic combination in (6) **impressed in*, as opposed to *impressed by* in standard English, is likely to be derived from *ปร้ะห้ะง* ‘impressed’ in Thai, which can be followed by *น้ะง* ‘in’. Cross-linguistic influence from Thai can also be noticed in (7), where *close* and *with* form a collocational deviation, while the appropriate preposition is *to*. In Thai, the word *ค้ะง* ‘close’ usually precedes the preposition *ค้ะง* ‘with’, which possibly explains the error in (7).

(8) * My home *stays at* Nakhon Si Thammarat.

Apart from the problems with prepositions as mentioned above, Thai EFL learners also appeared to be troubled with mother tongue interference when dealing with collocations involving nouns and verbs. As can be seen in (8), the deviant noun + verb collocation is clearly constructed based on Thai structure. That is to say, when a place or location is being referred to in Thai, the verb *ย้ะง* ‘stay or live’ and the preposition *ท้ะง* ‘at’ are used together as in *ย้ะง ท้ะง* ‘be located at/in’. Thai learners, especially those with limited English proficiency, may have a tendency to transfer such a verb phrase to English, leading to a deviant combination.

(9) * I *domesticate fishes* at home.

Meanwhile, certain incorrect verb + noun collocations are clearly linked to the learners’ first language. As in (9), even though the combination *domesticate fishes* is meaningful and grammatical, it probably does not reflect what the learner wanted to impart to the readers. In fact, the participants were asked to write about their hobbies, one of which could be keeping fish as a pet. In such a context regarding hobbies, the proper verb to co-occur with *fish* should be *have* or *keep*, as confirmed by the British National Corpus (BNC) data, whereas *domesticate* actually means “to bring animals or plants under human control in order to provide food, power or companionship” (Cambridge advanced learner’s dictionary, 2008, p. 417). This inappropriate verb may be literally translated from the verb *ล้ะง* ‘domesticate or raise’ in Thai.

(10) * We *play internet* in our free time at school.

(11) * She and I *play computer* if we have no homework to do.

The collocational errors in (10) and (11) may be attributed to L1 transfer as well. In Thai, it is possible to use to verb *ล้ะง* ‘play’ with the English loan words *อินเ้ะง* ‘the Internet’ and *ค้ะง* ‘computer’. Accordingly, Thai learners of English who are sometimes unaware of the mismatches between Thai and English could violate the collocational restrictions in L2 English by transferring *play internet* and *play computer* from their native tongue.

Apart from nouns and verbs, Thai learners were also found to commit L1-based errors on collocations concerning adjectives, as indicated in (12)-(13).

(12) * Last week, I was so busy staying for days in *my working room*.

(13) * *The dwelling place* of my uncle is in Korat.

Despite being understood, the two italicized combinations in (12) and (13) are considered non-targetlike as

they look foreign or different from what is really used in native speaker corpora, e.g. BNC. To be specific, the collocation *working room* in (12) may strike a native speaker as peculiar because standard English speakers should be more used to *office*. The deviant combination as such may be directly translated from *hǎwng thamɣaan* ‘office’ in Thai, where *hǎwng* means ‘room’ and *thamɣaan* means ‘work’. Similar to (12), *dwelling place* in (13) is also viewed as non-standard, probably ascribed to Thai influence. The deviation arises since the word *dwelling* alone refers to a place where people live (Longman dictionary of contemporary English, 2009, p. 529). The occurrence of *place* as a collocater of *dwelling* clearly causes redundancy.

The collected data also evidence L1 transfer concerning an adverb as part of an English collocation. As shown in (14), the adverb *unforgettably* and the modified verb *recalled* are obviously redundant. According to Longman dictionary of contemporary English (2009, p. 1450), the verb *recall* means “to remember a particular fact, event, or situation from the past”. Thus, the meaning of *unforgettably* is subsumed under *recall*, and the co-occurrence of these two words turn out to be awkward or unnatural in the target language. This particular problem probably arises from cross-linguistic influence from Thai, which allows the phrase like *jamdâi mǎjlym ‘recall’* (*jamdâi* ‘able to remember’ and *mǎjlym* ‘not forget’). Those learners, unaware of the redundancy in L2 resulting from both *unforgettably* and *recalled*, perhaps end up producing such a deviant combination.

(14) *I just *unforgettably recalled* the meaning a year ago.

To summarize, native language transfer seems to be the most common strategy which Thai students apply when dealing with collocation production in L2 English. The learners, by and large, are dependent on L1 equivalents when they fail to find the lexical items, i.e. nouns, verbs, adjectives, and adverbs, or grammatical items, e.g. preposition, in the L2. Their heavy reliance on L1 frequently leads to collocational errors in L2. It is also worth noting that not only are low-proficiency learners found to considerably transfer from L1 collocational knowledge, but also those with high-proficiency level appear to largely seek advice from L1 collocations. This provides support for Koya (2003), who surprisingly discovered that even Japanese EFL learners whose L2 proficiency is high tended to resort to L1 in selecting possible collocations for L2.

As can be seen in this section, the learners’ mother tongue perspicuously plays a significant role in affecting their L2 collocational learning. Put simply, most of the students’ collocational problems are attributed to L1 influence. All the collocational errors discussed in 5.1.1 are clearly interlingual. The findings of this current study are in line with those of a great number of past studies (e.g. Bahns, 1993; Bahns & Eldaw, 1993; Fan, 2009; Huang, 2001; Koya, 2003; Nesselhauf, 2003, 2005; Ying, 2009). Additionally, the present research project also corroborates previous studies on Thai EFL learners which reported on strong evidence of L1 influence (Boonyasquan, 2006; Mongkolchai, 2008).

In the next subsection, another learning strategy, known as synonymy, is fully discussed as a source of L2 English erroneous combinations.

5.1.2 Synonymy

The Thai learners in this current study were seen to apply synonymy in their acquisition of English collocations. In other words, they appeared to replace a word with another having a similar meaning. Although doing so may occasionally be successful, most studies indicate an obvious flaw of such a strategy (e.g. Boonyasquan, 2006; Farghal & Obiedat, 1995; Mongkolchai, 2008; Zughol & Abdul-Fattah, 2001).

According to the collocation data obtained, the Thai learners of English seem to employ synonymy as a strategy for learning L2 collocations. More precisely, they were found to use this particular strategy more in lexical collocations than in grammatical counterparts. (15)-(18) below are examples of lexical errors probably originating from synonymy.

(15) * I always prefer to stay in my *peaceable home* on the weekend.

(16) * I *authentically believe* that I have come too far.

(17) * *Some kinds of animals have to immigrate* for finding food.

(18) * *Home differentiates* with house.

The error in (115) should stem from the learners' confusion over the two semantically-related adjectives, *peaceful* and *peaceable*. While *peaceable*, usually seen in a literary text, is used either with a person who is not fond of fighting or a quarrel (Longman dictionary of contemporary English, 2009, pp. 1281-1282), the more frequent adjective *peaceful* is normally used with a time, place, or situation that is quiet and calm (Longman dictionary of contemporary English, 2009, pp. 1282). Additionally, the corpus-informed data from British National Corpus (BNC) demonstrate a number of tokens of *peaceful home*, but none of *peaceable home*. According to the two native English speakers as the informants, *peaceable home* strikes them as strange or unfamiliar, whereas *peaceful home* is absolutely acceptable. Therefore, the word *peaceful* is clearly more appropriate such a context as (15).

As regards (16), the source or error is apparently pertinent to synonymy as well. That is, even though the words *truly* and *authentically* are common in some aspects, both being associated with describing something real (Longman dictionary of contemporary English, 2009, p.17 & p. 1892), only *truly* can co-occur with the verb *believe*, as confirmed by the four collocation dictionaries, the BNC data, and the native speakers.

In (17), the deviant construction arises from the verb *immigrate* as a collocate of the noun *animals*. According to Longman dictionary of contemporary English (2009, p. 878), *immigrate* means "to come into a country in order to live there permanently". The BNC data also give further interesting details indicating that the subject nouns normally preceding *immigrate* are human, e.g. *people, spouses, ethnic Germans, his father and mother, outlaws, more married foreign gays, Indonesians, settlers, Indian nationals, residents*, etc. Thus, the verb *immigrate* should not be used with animals, as in (17). A more appropriate word choice should be *migrate*, which is generally related to how animals, e.g. birds, travel regularly from a part of the world to another (Longman dictionary of contemporary English, 2009, p. 1105). The BNC data also provide strong support for the dictionary information because a number of the collocates of *migrate* are animals, e.g. *swallows, the adult worms, birds of prey, salmon, animals, cranes, some fish, the larvae, the mammals, earthworms, whales, insects*, etc.

The collocational problem in (18) is also evidently caused by synonymy. Thai speakers may be confused over the two words *differ* and *differentiate*, both of which are close in meaning. The verb *differ* is different from *differentiate* in that the former means "to be different from something in some way" (Longman dictionary of contemporary English, 2009, p. 497), while the latter means "to recognise or express the difference between things or people" (Longman dictionary of contemporary English, 2009, p. 469). The distinction as such being made, the verb *differ* is needed here rather than *differentiate* since the learner probably wants to discuss some differences between *home* and *house*.

In addition to the use of synonymy strategy as a means of learning lexical collocations, this particular strategy was also discovered in grammatical collocations, as in (19).

(19) * We gave the present *for* our teacher.

In (19), the existence of the preposition *for* does not collocate with the verb *gave*. As a matter of fact, the verb *give* subcategorizes for an object noun phrase and a prepositional phrase beginning with *to*, as in *give something to someone* (Longman dictionary of contemporary English, 2009, p. 738). It is assumed that the mis-collocation (19) could be derived from the confusion over the proper use of the synonyms *give* and *provide*. Whereas *give* requires the preposition *to*, *provide* occurs in the construction *provide something for someone*. Such a difference in preposition as an essential element to constitute a grammatically correct structure for each verb may lead EFL learners to easily get confused about the usage of both verbs, possibly triggering collocational deviations in learner language.

(20) * They can provide you ___ plenty of time.

It is also found that the error in (20) may be caused by synonymy. In particular, the verb *provide* could probably be used by analogy with the verb *give*. That is, while *give* requires an indirect object and a direct object in that order, i.e. *give someone something* (Longman dictionary of contemporary English, 2009, p.

738), *provide* needs the preposition *with* between the two objects, i.e. *provide someone with something* (Longman dictionary of contemporary English, 2009, p. 1398). The student, producing the error in (20), might assume that *provide* can be employed in the same manner as *give*, thereby omitting the preposition *with*.

The findings with regard to synonymy as one of the strategies on which Thai EFL learners rely lend support to several previous research studies on second language acquisition of English collocations (e.g. Farghal & Obiedat (1995); Howarth, 1996, 1998; Zughol & Abdul-Fattah, 2001). The particular results are also consistent with the past studies on Thai learners (e.g. Boonyasquan, 2006; Mongkolchai, 2008), which found collocational errors induced by synonymy strategies.

The next subsection deals with another learning strategy, i.e. overgeneralization, also considered to be an actual source of deviant collocations.

5.1.3 Overgeneralization

Thai learners, in addition to the previously outlined strategies, also appear to make use of overgeneralization in the acquisition of English collocations. According to Richards, Schmitt, & Platt (2002), overgeneralization is a common process in L2 learning in which “a learner extends the use of a grammatical rule of linguistic item beyond its accepted uses” (p. 260). In the present research work, the participants presumably overgeneralized some words and form L2 miscollocations, as shown in the examples below.

(21) * The room that I *very like* is the dining room.

(22) * I *very love* to stay at home.

The collocational misuse in (21)-(22) may be attributed to an overgeneralization of the intensifier *very*. According to Longman dictionary of contemporary English (2009, p. 1949), *very* is regarded as an adverb used to emphasize an adjective, as in (23), where *very* modifies the adjective *cold*; an adverb, as in (24), where *very* modifies the adverb *slowly*; a phrase, as in (25), where *very* modifies the adjective phrase *important to that area*.

(23) It feels very cold today.

(24) The traffic’s moving very slowly this morning.

(25) The fishing industry is very important to the area.

(Longman dictionary of contemporary English, 2009, p. 1494)

Despite an adverb, the word *very* cannot modify a verb or verb phrase, which is why *very like* and *very love* in (21) and (22) respectively are considered ungrammatical. Nevertheless, Thai EFL learners whose proficiency is limited may be unaware of the restriction of *very*, thus improperly extending it to a verb or verb phrase modification.

Additionally, overgeneralization may be responsible for some deviations caused when a single word can occur in different parts of speech.

(26) * I not only *contacted with* old friends who is far away.

(27) * We don’t need to *use of* money to buy any gifts.

With regard to (26), the collocational problem occurs since the transitive verb *contact* actually requires a direct object rather than a prepositional phrase. The learner could be influenced by the usage of *contact* as a noun, where the preposition *with* is indispensable, as in (28). The presence of *with* is possibly overgeneralized to the verb *contact*, which results in an incorrect grammatical collocation in English.

(28) I haven’t had any contact with her for at least two years.

(Longman dictionary of contemporary English, 2009, p. 362)

Like (26), the combination of the verb *use* and the preposition *of* in (27) is collocationally erroneous probably because the learner got mixed up over the two word classes of *use*, i.e. a noun and a verb. When

use occurs as a noun, the preposition *of* often follows, as in (29).

(29) The increasing use of computers in education.

(Longman dictionary of contemporary English, 2009, p. 1937)

For this reason, the learner may extend the existence of the preposition *of* to the verb *use*, committing an error on English collocation. On the contrary, the verb *use* subcategorizes for a direct object instead of the prepositional phrase introduced by *of*, as in (30).

(30) I always use the same shampoo.

(Longman dictionary of contemporary English, 2009, p. 1937)

(31) * My kitchen is *on the back* of my house.

Furthermore, in (31), the prepositional phrase *on the back of my house* should be replaced by *at the back of my house* since the latter is generally used to refer to a place or location, as in (32).

(32) a small shop with an office at the back.

(Longman dictionary of contemporary English, 2009, p. 104)

It is assumed that the student may be more familiar with *on the back* or *on someone's back*, which is more frequent in the BNC data but does not fit here as it usually refers to the particular part of the body (Longman dictionary of contemporary English, 2009, p. 104), as in (33)-(34), or the part of something that is furthest from the front, as in (35).

(33) He lay on his back and gazed at the ceiling.

(34) The girl appeared again, now with a little baby on her back.

(35) He kissed her on the back of her head.

(Longman dictionary of contemporary English, 2009, p. 104)

The learner apparently overgeneralized *on the back* to a place or location, as seen in (35), producing an erroneous grammatical collocation.

The overgeneralization as the learners' strategy revealed by the current research project is commonplace in the process of L2 acquisition of collocations, as confirmed by previous studies (e.g. Shih, 2000; Zughol & Abdul-Fattah, 2001).

6. Conclusion

As indicated in the findings of the present study, Thai EFL learners no matter how L2-proficient they are really have collocational problems. In particular, according to both high-and-low proficiency learners' data, the preponderance of errors in lexical collocations over those in grammatical ones is prominent. With respect to the origins of these deviations, Thai learners' errors seem to be related to certain learning strategies. In a nutshell, the learners' use of collocations is strongly influenced by their L1 collocational knowledge. In addition to native language transfer, they are also inclined to depend upon synonymy and overgeneralization, which are evidently regarded as major causes of erroneous collocational production.

7. Pedagogical implications

Like other EFL learners, Thai students clearly use some similar deviant English collocations in their writing, as discussed earlier. To be more precise, the common strategies, such as synonymy and overgeneralization, appear to be universally shared by EFL learners in general in their second language acquisition of English collocations, and these lead to certain erroneous combinations in L2 English. Not only English teachers in Thailand, consequently, but also those instructors elsewhere whose students are in an EFL setting may, to a great extent, benefit from this research study since their students are supposed to be confronted with the similar problems resulting from an application of such strategies. That is, they can prepare their collocation lessons which substantially focus on the genuine problems facing EFL learners. Teachers may create collocation exercises based on the common combinations that are obviously problematic. Moreover, as suggested by this study, the learners' native language has a strong, significant impact on their use of English

collocations. It is recommended that Thai EFL teachers highlight the transfer-based problems caused by the differences between English and Thai, with a particular emphasis on the deviations frequently occurring in their interlanguage (Nesselhauf, 2005). Hopefully, the number of serious collocational errors should diminish through such an instruction underscoring those L1-related problems as well as developmental errors triggered by other learning strategies.

References

- Bahns, J. (1993). Lexical collocations: A contrastive view. *ELT Journal*, 47(1), 56-63.
- Bahns, J. & Eldaw, M. (1993). Should we teach EFL students collocations? *System*, 1(1), 101-114.
- Benson, M., Benson, E., & Ison, R. (1997). *The BBI dictionary of English word combinations*. Amsterdam: John Benjamins.
- Bley-Vroman, R. (1983). The comparative fallacy in interlanguage studies: The case of systematicity. *Language Learning*, 33, 1-17.
- Boonyasaquan, S. (2006). An analysis of collocational violations in translation. *Journal of Humanities*, 27, 79-91. Bangkok: Faculty of Humanity, Srinakharinwirot University.
- Cambridge advanced learner's dictionary* (2008). Cambridge : Cambridge University Press.
- Chomsky, N. (1989). A Review of B. F. Skinner's Verbal Behavior. *Language*, 35, 26-58.
- Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics in Language Teaching*, 5, 161-170.
- Ellis, R. (1985). *Understanding second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2008). *The study of second language acquisition*. Oxford: Oxford University Press.
- Fan, M. (2009). An exploratory study of collocational use by ESL students—A task based approach. *System*, 37, 110-123.
- Farghal, M. & Obiedat, H. (1995). Collocations: A neglected variable in EFL. *International Review of Applied Linguistics in Language teaching*, 33(4), 315-531.
- Firth, J. R. (1957). *Papers in linguistics, 1934-1951*. Oxford: Oxford University Press.
- Gass, S. & Selinker, L. (2008). *Second language acquisition: An introductory course*. New York: Routledge.
- Granger, S. (1998) Prefabricated patterns in advanced EFL writing: Collocations and formulae. In Cowie, A. (ed.) *Phraseology: Theory, analysis and applications* (pp.145-160). Oxford: Oxford University Press.
- Granger, S., Paquot, M., and Rayson, P. (2006). Extraction of multi-word units from EFL and native English corpora. The phraseology of the verb 'make'. In A. Hacki Buhofer and H. Burger (Eds.), *Phraseology in motion I: Methoden und Kritik*. (Basel, 2004) (pp. 57-68). Baltmannsweiler: Schneider Verlag Hohengehren.
- Hill, J. (2000). Revisiting priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), *Teaching collocation: Further development in the lexical approach* (pp. 47-69). London: Commercial Colour Press Plc.
- Hill, J., & Lewis, M. (1997). *LTP dictionary of selected collocations*. Hove: Language Teaching Publications.
- Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics*, 19, 45-72.
- Huang, L. (2001). Knowledge of English collocations: An analysis of Taiwanese EFL learners. [Online] Available: www.utexas.edu/students/flesa/tpfle/contents7.doc
- Koya, T. (2003). A study of collocation in English and Japanese noun-verb combinations. *Intercultural Communication Studies*, XII-1, 125-145.

- Lado, R. (1957). *Linguistics across culture*. Ann Arbor, MI: University of Michigan Press.
- Langan, J. (2009). *College writing skills*. New York: The McGraw-Hill Companies.
- Lewis, M. (1997). *Implementing the lexical approach*. Hove, English: Language Teaching Publications.
- Lewis, M. (2000). *Teaching collocation: Further development in the lexical approach*. London: Commercial Colour Press Plc.
- Longman dictionary of contemporary English 5th edition* (2009). Essex: Pearson Education Ltd.
- Mclaughlin, (1987). *Theories of second language learning*. London: Arnold.
- Mongkolchai, A. (2008). *A study of university students' ability in using English collocations*. Master's project, Srinakharinwirot University.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, 24(2), 223-242.
- Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: John Benjamins.
- Oxford collocations dictionary 2nd edition*. (2009). Oxford: Oxford University Press.
- Phoocharoensil, S. (2010). A corpus-based study of English synonyms. *International Journal of Arts and Sciences*, 3(10), 227-245.
- Rundell, M. (2010). *Macmillan collocations dictionary for learners of English*.
- Odlin, T. (1989). *Language transfer*. Cambridge: Cambridge University Press. Oxford: Macmillan Publishers Limited.
- Richards, J.C., Schmidt, R. & Platt, H. (2002). *Longman dictionary of language teaching and applied linguistics*. London: Longman.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, 10, 209-231.
- Selinker, L. (1992). *Rediscovering interlanguage*. London: Longman.
- Shih, R. (2000). Collocation deficiency in a learner corpus of English from an overuse perspective. *Proceedings of 14th Pacific Asia Conference on Language, Information and Computation*, 281-288.
- Sinclair, J. (1990). *Corpus concordance collocation*. Oxford: Oxford University Press.
- Swan, M. (2005). *Practical English usage*. Oxford: Oxford University Press.
- Tarone, E. (2006). Interlanguage. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics (2nd ed.)*, 5, (pp. 747-752). Oxford: Elsevier.
- Ying, S. (2009). Study on collocations in English writing by Chinese students. *Sino-US English Teaching*, 6(3), 25-30.
- Zughoul, M. R. & Abdul-Fattah, H. S. (2001). Collocational competence of Arabic speaking learners of English: A study in lexical semantics. [Online] Available: http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1b/4f/28.pdf

Notes

Note 1. For some other researchers, various alternative terms have been used to refer to the same phenomenon, e.g. *approximative systems* (Nemser, 1971), *idiosyncratic dialects* and *transitional competence* (Corder, 1971).

Note 2. The students from Foundation English Course 1 were excluded from the present study since their English collocational competence could be so extremely low that they may not be capable of a variety of L2 English collocation production.

Table 1. Distribution of collocational deviations in the high-proficiency learners' interlanguage

Types of collocation		Number of deviant collocations	%
Lexical	verb + noun	22	25.28
	adjective + noun	12	13.79
	verb + adverb/ adverb + verb	6	6.90
	noun + verb	5	5.75
Total		45	51.72
Grammatical	verb + preposition	16	18.39
	preposition + noun	16	18.39
	adjective + preposition	7	8.05
	noun + preposition	3	3.45
Total		42	48.28
Total number of deviations		87	100

Table 2. Distribution of collocational deviations in the low-proficiency learners' interlanguage

Types of collocation		Number of deviant collocations	%
Lexical	verb + noun	45	40.54
	noun + verb	13	11.71
	adjective + noun	6	5.41
	verb + adverb	1	0.90
Total		65	58.56
Grammatical	verb + preposition	26	23.42
	preposition + noun	14	12.61
	adjective + preposition	6	5.41
	noun + preposition	-	-
Total		46	41.44
Total number of deviations		87	100