Thai EFL Learners' Knowledge of Congruent and Incongruent Academic L2 Collocations

Waraphorn Khantiwong[a], Atikhom Thienthong[b,*]

[a] waraphorndr@gmail.com, English and Communication Program, Ubon Ratchathani University, Thailand
[b] atikhom.k@ubu.ac.th, English and Communication Program, Ubon Ratchathani University, Thailand
[*] Corresponding author, atikhom.k@ubu.ac.th

Abstract

This article reports the results of a study that explored Thai EFL learners' repertoire of congruent and incongruent academic English (L2) collocations in relation to their native-Thai language (L1) and academic experience. Eighty Thai tertiary students performed a gap-filling translation test on 15 congruent and 15 incongruent collocations by providing equivalent L2 collocates. The results indicated that the students' greater exposure to academic discourse increased their acquisition of academic collocations. However, overall they demonstrated insufficient knowledge of academic collocations, especially incongruent L1-L2 combinations. This is because they depended heavily on their L1 lexicon and general L2 lexis to compensate for their lack of knowledge and awareness of typical academic L2 collocations. The results suggest that although L2 exposure has a facilitative effect on collocation acquisition, it is still imperative that EFL learners receive explicit instruction which is devoted to elaborating and disambiguating the L1 and L2 meanings of

	academic collocations and their component words in isolation.

1. Introduction

Collocation, a typical dimension of formulaic language, is important for language learning and fluent language use, given that units of meaning are determined by phrases, not words (Sinclair, 2008). With regard to language acquisition, learners learn more efficiently as they are exposed to larger multiword units rather than discrete single-word instances of language. As a result, when they acquire an increased knowledge of collocation, they can process and produce formulaic expressions more effortlessly and fluently than non-formulaic ones (Conklin & Schmitt, 2012). While learners' ability to use formulaic expressions is considered a key indicator of native-like and advanced proficiency, combining words which typically appear together as a meaningful chunk is cognitively challenging for learners whose native language is not English (Snoder & Reynolds, 2019). This challenge becomes even greater for them with respect to academic collocations which are more characteristic of specialized than general discourses.

Collocation has increasingly received considerable attention in L2 research on formulaic language. Over the past decade, there have been a growing number of empirical studies into learners' productive knowledge of collocations, such as adjective-noun collocations in essay writing (Siyanova-Chanturia & Schmitt, 2008), verb-noun collocations in L2 writing (Laufer & Waldman, 2011), adjective-noun collocations in the writing of beginner learners (Siyanova-Chanturia, 2015), verb-noun collocations in native and non-native student writing (Yoon, 2016), collocations in international business management discourse (Chansopha, 2018), and acquisition of verb-noun collocations in explicit instruction (Tsai, 2018). These studies are similar in that they each focused on one syntactic pattern and general collocations. Other research has investigated learners' acquisition of collocations with respect to frequent exposure and collocation distance, including adult retention of collocations (Durrant & Schmitt, 2010), learner intuition of collocation frequency (Siyanova-Chanturia & Spina, 2015), effect of frequency and exposure on collocational knowledge (Fernández & Schmitt, 2015), and processing of adjacent and non-adjacent collocations (Vilkaitė & Schmitt, 2019). These

studies indicate that adequate exposure and language proficiency predict learners' knowledge of collocations.

Many investigations have been conducted to explore the L1 influence on nonnative-English learners' acquisition of congruent and incongruent collocations. The results of these investigations seem to point in a similar direction. Yamashita and Jiang (2010) reported that Japanese ESL learners generally outperformed their Japanese EFL counterparts, but the two groups made more incorrect choices of incongruent than congruent collocations. A few years later, a study by Wolter and Gyllstad (2013) showed that advanced L2 learners (L1 Swedish) had significantly more difficulty with incongruent than congruent collocations. In addition, Peters (2015) found that incongruent collocations placed a greater learning burden on EFL learners (L1 Dutch) as they combined single words based on their L1 concepts. In the Thai context, research on collocation congruency is relatively scare, except for two research studies by Phoocharoensil (2013) and Sanguannam (2017). Phoocharoensil (2013) studied Thai EFL students' productive use of collocations in descriptive essays and found that generally the students had problems with incongruent collocations as evidenced by their heavy reliance on their L1 lexicon. Similarly, Sanguannam (2017) found that Thai EFL students at both intermediate and advanced levels made significantly more errors on incongruent collocations than congruent ones. However, none of the studies reviewed above investigated EFL learners' knowledge of congruent and incongruent L1-L2 academic collocations which are inherently different from general collocations in terms of semantic transparency and combination restrictiveness.

Previous research has shown that collocational knowledge is associated with language status and learning experience. Generally, native-English speakers who are more linguistically competent than non-native learners have a large number of firmly stored formulaic sequences at their disposal. However, attaining such a level of collocational competence is a cognitive challenge for EFL learners. This is largely because there are many collocations whose component words are not semantically equivalent in the L1 and L2. Consequently, if EFL learners produce collocations based on their L1 lexicon, the resulting combinations are likely to be uncommon or infelicitous. This semantic phenomenon is known as L1-L2 incongruency (Snoder & Reynolds, 2019; Wolter & Gyllstad, 2013). Because of such inequivalence in meaning, adequate

exposure is seen as having a significant effect on EFL learners' acquisition of collocations. This effect can also be observed in academic collocations. Evidence indicates that students' greater exposure to academic discourse promotes their collocational knowledge, and this academic experience is more efficient in acquiring collocations than native-English-speaker status (Frankenberg-Garcia, 2018). Although learners' mother tongue and L2 experience are subject to learning collocations, whether or not EFL learners' L1 and academic experience influence their knowledge of congruent and incongruent collocations in academic discourse remains relatively underexplored.

2. Literature Review

2.1 Academic collocation

Collocation has no universal definition because it can be approached by different theoretical orientations, namely, psychological, phraseological, and statistical (Hoey, 2005; Men, 2018; Nesselhauf, 2005). The present study draws on the phraseological and statistical notions which approach collocations from their semantic properties and frequent co-occurrence of component words. A phraseological approach defines collocation as a semantic combination configured by relative transparency in meaning and restrictively collocating words (Laufer & Waldman, 2011). According to a statistical approach informed by corpus linguistics, collocation is defined as the frequent co-occurrence of two lexical words, that is, node and collocate (Stubbs, 2009). We use these two definitions for two reasons. First, we intend to explore L1 vis-à-vis L2 collocational meanings. Second, we focus on common collocations. Hence, it is essential that the collocations in question occur frequently in a salient fashion so that EFL learners are likely to encounter them incidentally in their learning discourse.

The co-existence of a node and its collocate involves not only semantic but also syntactic relations. Syntactically, words do not occur in arbitrary sequences, but they collocate with nearby words to form meaningful collocations. As such, collocations are highly patterned; their elements co-occur in context because they are syntactically attracted (Flowerdew, 2012). In their co-occurrence, the nodes and collocates are associated in different syntactic patterns, such as adjective + noun

(*important role*), verb + noun (*conduct research*), adverb + verb (*strongly suggest*), and adverb + adjective (*significantly different*). The present study targets collocations with these syntactic structures because they are characteristic of academic discourse (Lei & Liu, 2018). Moreover, they pose challenges for EFL learners who have limited language proficiency and academic experience. These challenges occur largely because many collocations are semantically unpredictable. That is, the actual senses of their component words are dependent on their typical collocates and syntactic properties, such as *suggest, conduct,* and *strongly*. These forms of linguistic behavior should be made explicit to learners.

It is generally recognized that language use is different between academic and non-academic registers. With respect to collocation, academic and non-academic discourses are distinctive in several semantic and syntactic properties. In general contexts, many collocations are typically characterized by delexical verbs, such as *take a look*, *have a problem*, and *make a mistake* (Wang, 2016). Delexical verbs (e.g. *take*, *have*, *make*, *give*, *get*, and *do*) are light verbs that express basic, yet little meaning and require a noun as their syntactic argument (Altenberg & Granger, 2001; Huddleston & Pullum, 2017). This means that delexical verbs always co-occur with other nouns, known as delexical collocations. Semantically, however, they are not totally compositional (Nagy T. et al., 2020); their meaning is mainly derived from their collocates (*look*, *problem*, *mistake*) rather than their light nodes (*take*, *have*, *make*) (Stubbs, 2001). In other words, the delexical node is semantically equivalent to the collocate (Huddleston & Pullum, 2017). In terms of collocability, the nodes of delexical collocations tend to occur freely with many other collocates (Liu & Lei, 2018; Stubbs, 2001). For example, the node *have* collocates with, such as, *fun*, *chance*, *access*, *power*, *problem*, *accident*, *difficulty*, *trouble*, *plan*, and *effect*.

In academic contexts, by contrast, delexical collocations are not common (e.g. Lei & Liu, 2018). It is known that academic discourse is lexically dense and highly phrasal as it is structurally compressed, with lexical subjects and pre-and post-modifying phrases (Biber & Gray, 2010). These defining characteristics render academic collocations markedly different from general collocations in terms of semantic and syntactic properties. Semantically, academic collocations are compositional; their components (i.e. node and collocate) convey their respective meanings to constitute a meaningful combination. They are combined in different

syntactic structures due to their structural compression. Another typical feature of academic discourse is that academic collocations are relatively restricted combinations; they keep particular company which is rarely substitutable (e.g. Lei & Liu, 2018). For instance, when *strong* typically collocates with *evidence*, it means '*convincing*'. In this sense, *strong* cannot be replaced by its synonym *powerful* and prefers *evidence* instead of *proof*. The two components *strong* and *evidence* usually co-occur and contribute their own meanings to the entire collocation.

2.2 Congruent and incongruent collocations

Learners' L1 plays an important part in their L2 acquisition. This is why it is very challenging for native-Thai learners to master English. It is even challenging for them to acquire felicitous collocations. While many L2 collocations have an equivalent L1 meaning, others do not have such semantic equivalence. These interlingually semantic phenomena are described as L1-L2 congruency and incongruency, respectively. Congruent collocations refer to combinations whose components are semantically similar or identical in the L1 and L2 while incongruent collocations refer to combinations whose components are semantically different in the L1 and L2 (Yamashita & Jiang, 2010). For EFL learners, incongruent collocations are more difficult to learn and acquire than congruent collocations. This is because while learners can produce congruent collocations correctly depending on their L1 lexical senses and concepts, doing so with incongruent collocations can cause erroneous collocation choices (Snoder & Reynolds, 2019; Webb & Kagimoto, 2011). Many collocations are incongruent because their component words are semantically and conceptually idiosyncratic from their combinations, such as *play* in *play role*, *poor* in *poor performance*, and *strongly* in *strongly linked*.

Although it is a challenge to teach and learn collocations, some proposals have been suggested. Empirical studies indicate that EFL learners can increase their collocational knowledge incidentally through implicit methods as they are frequently exposed to target collocations. In academic contexts, they can also benefit from such indirect processes of acquiring academic collocations as they engage in language use and learning activities during university years. This acculturation promotes learners' frequent exposure to academic collocations (Frankenberg-Garcia, 2018). However, research indicates that it is not adequate for EFL

learners to implicitly acquire incongruent L2 collocations which cannot be combined literally depending on their L1 lexical knowledge (Wolter & Gyllstad, 2013). Semantically, moreover, many incongruent collocations are unpredictable from their component words in separation which learners may have initially acquired and established in their L2 lexical network. Hence, incongruent collocations deserve special attention in explicit learning methods. Webb and Kagimoto (2011) suggest that teachers should assist EFL students by raising their awareness of the L1 and L2 forms and meanings of incongruent collocations with a low degree of translational equivalence.

Although L2 exposure has been acknowledged as promoting learners' implicit acquisition of general collocations in other EFL contexts, there have been relatively few studies to explore EFL learners' implicit acquisition of academic collocations, especially congruent and incongruent ones. In Thai contexts, collocation tends to receive peripheral attention compared to vocabulary in L2 pedagogy and research. One plausible explanation for this is that Thai EFL teachers concentrate on teaching students single words because they usually assume that students can automatically acquire collocations by assembling the words they have learnt individually. This collocation learning process is common among L2 classroom learners (Wray, 2002). To promote the value of formulaic language, therefore, this study explores Thai EFL learners' knowledge of congruent and incongruent collocations they are likely to encounter in academic discourse. We hope that this study can provide some important insights into Thai learners' knowledge of academic collocations. Also, it can reveal whether and which collocations (congruent or incongruent) are treacherous for Thai learners to acquire. The results suggest some pedagogical implications for implementing instructional interventions to maximize academic collocation learning.

3. Research Methods

3.1 Research questions

This study aims to explore Thai tertiary students' knowledge of academic collocations between two lexical items which are semantically congruent and incongruent in their L1 and L2. In accordance with this aim, the study answers the following questions:

1. To what extent does Thai EFL learners' exposure to academic discourse influence their knowledge of academic collocations?

2. Does Thai EFL learners' native-language status influence their knowledge of academic collocations?

3. Do the number and quality of academic collocations vary according to Thai EFL learners' academic experience?

3.2 Participants

The study involved a total of 80 Thai undergraduate students. At the time of investigation, they were studying for a major in English and Communication in a four-year Bachelor's degree program at a public university in North-eastern Thailand. The students' exposure to English had been limited to courses which did not explicitly teach them academic collocations. In this study, we used a non-proportionate stratified sampling method to reach an equal number of samples from each university year for a comparable investigation. Thus, we recruited 20 students from four different university years ($n = 80$) who were divided into two groups of 40 participants according to their academic learning experience. The first- and second-year students were labelled as the less-experienced group (LE) ($n = 40$) while the third- and fourth-year students were regarded as the more-experienced group (ME) ($n = 40$). The division into two groups allowed us to explore their collocational knowledge in relation to their exposure to academic discourse. Due to their different years of study at university, they were considered relatively heterogeneous in terms of linguistic competence and academic experience.

3.3 Collocation test

The instrument was a gap-filling translation test. It consisted of 15 congruent collocations and 15 incongruent collocations. Each of the collocation sets had four syntactic patterns with the collocates marked in **bold**: adjective + noun (***strong*** *evidence*), verb + noun (***receive*** *attention*), adverb + verb (***strongly*** *suggest*), and adverb + adjective (***purely*** *academic*). The tested nodes and collocates (see Appendices 1 and 2) were typically found across academic texts and registers because they were selected from general academic words compiled by Gardner and Davies (2014) and academic collocations examined by Lei and Liu (2018) in general academic

corpora and sub-corpora, such as the British Academic Written English Corpus (BAWE), the British National Corpus (BNC), and the Corpus of Contemporary American English (COCA). The collocations measured were therefore deemed suitable to investigate Thai EFL students' knowledge of academic collocations.

To make sure that the selected nodes and collocates typically and frequently co-occurred, we determined their association strength and co-occurrence frequency (*F*) favored in the enTenTen15 corpus. The enTenTen15 corpus, accessed via Sketch Engine, is a web corpus of 15 billion words compiled from a wide range of written texts. The association strength is measured by logDice (*L*) in Word Sketch, a feature of Sketch Engine. LogDice is a statistic measure of collocation typicality, which is sensitive to meaningful and low-frequency collocations. It operates on a certain scale with a maximum of 14 where high values indicate significant collocations (Rychlý, 2008). The collocations with a logDice score of 5-14 and a minimum of 1,000 co-occurrences were selected and included in the collocation test (see Appendices 1 and 2). These criteria were adopted to ensure that the target collocations were meaningful and common.

The two corpus-based criteria were applied to selecting both congruent and incongruent collocations. In this study, congruent L2 collocations are lexical combinations whose node and collocate senses are semantically equivalent to the L1. For example, *receive attention* is a congruent collocation where the collocate *receive* means '*dâj ráp*' and the node *attention* '*khwaːm-sǒn-caj*'. In contrast, incongruent L2 collocations are lexical combinations whose collocate senses are not semantically equivalent to the L1. For example, *play role* is considered incongruent. Based on the L2 sense, the node *role* means '*bòt-bàːt*' and the collocate *play* '*miː*'. But if *miː* is literally translated, its L2 translation equivalent is *have*. The collocate *have* is not incorrect, but it is a delexical word and not as typical as *play*. If the participants were not aware of the node *role* and relied on their L1 lexical sense, they were very likely to supply *have.*

Each question item contained a sentence in the L2 with a target collocation whose collocate was removed and needed to be filled in by the participants. The lexical category of the removed collocate was also given after the sentence to avoid grammatically undesirable answers. The L2 sentence was fully translated into the L1 with only the equivalent of the collocate underlined and boldfaced to direct the participants' attention to the target L1 word. The L1 equivalent cued L2 collocates to be supplied by

the participants. In the following examples, the L2 and L1 sentences are illustrated:

Congruent collocation
Teachers need more financial support to _____ research. (verb)

kham-plɛː : khruː thôːŋkaːn kaːn-saʔ-nàp-sàʔ nǔn thaːŋ kaːn ŋən mâːk khɯ̂n phɯ̂a **tham** wíʔ-caj

kham-tɔ̀ːp :_____

Incongruent collocation
Excessive smartphone use can cause _____ performance in school. (adjective)

kham-plɛː : kaːn cháj sàʔ-máːt-foːn mâːk kəːn paj pen hèːt hâj phǒn-kaːn-rian **tòk-tàm**

kham-tɔ̀ːp :_____

The L2 sentences were taken from the BNC and COCA academic sub-corpora. They were modified for simplicity and validated by three native-English speakers for clarity. Then we translated the validated sentences word for word into the L1 and carefully checked the most typical translations of the target nodes and collocates by consulting L1-L2 bilingual dictionaries and the thTenTen18 (Thai) corpus available in Sketch Engine. Next, the translated sentences were proofread and verified by three teachers of Thai with L2 proficiency. This procedure was to ensure that the L1 translations in question were semantically equivalent to the L2 collocations.

3.4 Data collection and analysis

The study used a paper format test to assess the participants' productive knowledge of collocations. The test was distributed to the participants when they did not have a class. Before starting the test, they were assured that the test was entirely anonymous. The participants had 30 minutes to complete the test. The test instruction was written in the participants' L1 (i.e. Thai). The participants were not allowed to consult any reference resources. In the productive test, they were instructed to read each L2 sentence and provide a typical L2 collocate which was equivalent

in meaning to the L1 and co-textually appropriate. They wrote their answers in the given space. They were also allowed to supply multiple collocates. The responses did not require any grammatical inflections.

The responses provided by the participants were evaluated for their salient and strong collocations as follows. First, we filtered incorrect responses with different L1 and/or L2 senses from possible responses with similar L1 and L2 senses. For example, the L1 word *phóp* has several L2 translations, such as *find*, *meet*, and *encounter*. But given the context and node *problem* as in (1), only *encounter* was marked as correct. Second, we adopted two corpus-based measures: co-occurrence frequency and collocation typicality to judge the salient usage and strong association of possible responses (Rychlý, 2008). The collocations that have a logDice threshold of ≥ 5 and a co-occurrence of at least 500 times are acceptable. For example, the L1 word *tham* has three L2 translations: *do*, *conduct*, and *carry out*. These collocates are possible, given the context and node *research* as in (2). The most typical is *conduct* (F = 211,884, L = 10.98), followed by *do* (F = 176,910, L = 9.01) and *carry out* (F = 9,536, L = 6.83). The correct and incorrect responses were calculated by descriptive statistics and a chi-square test to determine significant differences.

(1) If you _____ any problems with our software, please contact microcompany@software.co.th.
(2) Teachers need more financial support to _____ research.

4. Results and Discussion

The participants provided a total of 2,788 collocate tokens. Of these tokens, 111 (3.98%) were excluded from the analysis because they were erroneous in terms of lexical classes and syntax. As a result, 2,677 tokens were included in the analysis.

4.1 Knowledge of academic collocations

Table 1 presents the sums of frequencies and percentages of collocable and non-collocable responses between the LE and ME participants. The results clearly indicate that the higher percentage of collocable responses was found in the ME group, accounting for 47.87%,

whereas the higher percentage of non-collocable responses was observed in the LE group, constituting 64.47%.

Table 1

Comparison of collocate tokens by academic experience

Participants	Collocability			χ^2	*p*
	Yes (%)	No (%)	Total (%)		
LE	459 (35.53)	833 (64.47)	1,292 (48.26)	41.83	0.00*
ME	663 (47.87)	722 (52.13)	1,385 (51.74)		
Total (%)	1,122 (41.91)	1,555 (58.09)	2,677 (100.00)		

*p < .01.

To answer the first research question, we compared the collocable and non-collocable responses provided by the LE and ME groups by using a chi-square test. The results showed that there was a significant difference between the two groups in the distribution of the correct and incorrect answers (χ^2 = 41.83, *df* = 1, *p* < .01). This shows that L2 learners who are extensively exposed to academic discourse are more likely to acquire a greater number of academic collocations. This evidence lends empirical support to many empirical studies confirming that exposure promotes collocation acquisition (Conklin & Schmitt, 2012; Fernández & Schmitt, 2015; Siyanova-Chanturia & Spina, 2015). However, although the ME students performed significantly better than their LE counterparts, the two groups' knowledge of academic collocations was relatively limited, as indicated by 58.09% of incorrect responses. This implies a need for explicit teaching instruction to boost academic collocation learning.

The results showed that the ME group outperformed the LE group on the majority of questions. The former tended to choose corpus high-frequency and typical collocations. The notable differences were observed in two typical collocations: *provide benefit* (*F* = 73,288, *L* = 7.32) and *conduct research* (*F* = 211,884, *L* = 10.98). These typical collocations were given by the ME students while the LE students supplied *give benefit* (*F* = 9,659, *L* = 4.94) and *do research* (*F* = 176,910, *L* = 9.01). According to the designated criteria, *give benefit* was marked as incorrect while *do research* was acceptable, yet it is useful to note that *do* is a delexical verb which tends to appear arbitrarily with other words in various registers. The result indicates that learners who have greater L2 exposure tend to produce

more common and meaningful collocations. This collocational knowledge is probably correlated with corpus frequency (Fernández & Schmitt, 2015; Siyanova-Chanturia & Spina, 2015). The correlation can be explained by the fact that when learners are frequently exposed to lexical items that typically co-occur, they recognize those previously encountered components as meaningful collocational chunks.

The overall results, however, demonstrated the participants' insufficient knowledge of academic collocations. While previous literature has confirmed that general collocations are challenging for EFL learners (Nesselhauf, 2003; Peters, 2015; Snoder & Reynolds, 2019; Yamashita & Jiang, 2010), the present study provides further evidence regarding L2 collocation research by showing that academic collocations are also problematic for EFL learners. Tertiary EFL learners seem to study English as "less a language than a basic academic skill" (Hyland, 2013, p.54). Arguably, they are more exposed to academic English than general English. Despite their greater exposure to formal academic discourse, they still have inadequate knowledge of academic collocations. Regarding the results in response to the first research question, there are some practical and linguistic reasons which explain EFL learners' limited knowledge of academic collocations. In L2 instruction, it is evident that collocation does not receive as much practical interest as single-word vocabulary in EFL contexts (e.g. Nguyen & Webb, 2016). Therefore, students may not be aware that in certain contexts, words semantically interact with other nearby words, and their actual meanings are largely determined by their collocates (Sinclair, 2008; Stubbs, 2009) and their register (Adamska-Sałaciak, 2013). For example, *give* and *provide* share a similar sense in Thai, that is, '*hâj*', but *provide* is a better choice because it typically collocates with the node *benefit*. But 54 participants chose *give* while 28 of them favored *provide*. Among the latter group, the majority (21 students) were the ME participants.

4.2 L1 influence on collocational knowledge

Table 2 summarizes the statistics for the two sets of academic collocations: congruent and incongruent. It can be seen that the numbers of collocate tokens supplied for the two collocation sets are relatively similar, with 50.73% found in the congruent collocations and 49.27% in the incongruent collocations.

Table 2

Comparison of congruent and incongruent collocations

Collocation	Collocability			χ^2	p
	Yes (%)	No (%)	Total (%)		
Congruent	627 (46.17)	731 (53.83)	1,358 (50.73)	20.53	0.00*
Incongruent	495 (37.53)	824 (62.47)	1,319 (49.27)		
Total (%)	1,122 (41.91)	1,555 (58.09)	2,677 (100.00)		

*$p < .01$.

In answer to the second research question, the results revealed a significant difference between the acceptable responses and the types of collocations ($\chi^2 = 20.53$, $df = 1$, $p < .01$). That is, while the participants from the two groups were relatively more successful in providing correct responses to congruent collocations, they generated significantly more incorrect answers to incongruent collocations. Of 1,319 tokens identified in the incongruent collocation set, only 495 (37.53%) were observed to be acceptable L2 collocates. This figure is much lower than the percentage of correct answers provided for congruent collocations. The results indicate that incongruent collocations were more challenging for both LE and ME groups. Furthermore, it is evident that those who had less L2 learning experience in academic contexts experienced greater difficulties.

The results clearly showed that the ME and LE participants were able to provide typical L2 collocates more successfully for congruent than incongruent collocations. For example, most of them correctly chose *widely accept* ($F = 25,944$, $L = 9.2$). The congruent L2 collocate *widely* is semantically equivalent to '*jàːŋ kwâːŋ-khwǎːŋ*'. This indicates that forming congruent collocations is more cognitively effortless for learners when they have already acquired L2 equivalents in their lexical structure (Wolter & Yamashita, 2017). However, such processing is not true of incongruent collocations. Our results revealed that the participants made significantly more incorrect choices of incongruent collocations. In many questions, they were unable to supply typical L2 collocates. For example, *poor performance* ($F = 24,383$, $L = 8.04$) which is equivalent in meaning to '*phǒn-kaːn-rian tòk-tàm*' is the significant collocation, but none of the participants supplied *poor*. In the given context, *poor* means '*tòk-tàm*'. If *tòk-tàm* is literally translated, it is equivalent to *low* which was given by 52 participants. For most Thai students, *poor* means '*having little money*'. In addition to *poor*

performance, other incongruent collocations shared similar results, such as *play role*, *raise awareness, encounter problem, careful analysis*, *serious problem*, *strong evidence*, *strongly linked*, and *purely academic.* The results suggest that incongruent L2 collocations combined from their polysemous collocates also causes trouble for EFL learners.

The present study found that the participants processed congruent collocations much more successfully than incongruent collocations. The findings suggest that their linguistic background negatively influenced their knowledge of incongruent collocations. These findings are similar to those of Peters (2015), Phoocharoensil (2013), Sanguannam (2017), and Wolter and Yamashita (2017). There are some plausible reasons to explain why congruent collocations are easier to learn than incongruent collocations. Yamashita and Jiang (2010) note that congruent collocations can be formed word for word effortlessly depending on their L1 equivalents. This L1-L2 congruency can greatly accelerate their processing of L2 collocations (Wolter & Yamashita, 2017). Thus, frequent L2 exposure may not be as important for congruent collocations as for incongruent collocations. This suggests that ESL/EFL learners tend to acquire congruent collocations easily and quickly in earlier stages of learning L2 collocations because there is a correspondence between the L1 and L2 (Wolter & Gyllstad, 2013). However, we argue that adequate input is still necessary for learners to acquire felicitous congruent collocations, given that many L1 lexical items can be translated into multiple L2 equivalents with similar or identical senses. These synonymous senses need to be disambiguated by syntax, register, and collocation. For instance, the participants supplied *earn*, *gain*, *get*, *have*, *obtain*, and *receive* which they translated from the L1 equivalent *dâj ráp*. Considering the given L2 node *attention, receive* and *gain* are the most appropriate for the context and register. If the participants were adequately exposed to *receive attention*, they would be likely to choose *receive* over other sense-sharing candidates in their L2 lexicon.

The fact that incongruent collocations cause difficulty for EFL learners can be explained by three reasons. First, the participants relied heavily on word-for-word translation equivalents in their L1 which were not semantically congruent with the L2 collocations. The participants' over-reliance on their L1 lexicon resulted in unusual L2 collocations which they solved by simply drawing on their L1 lexical knowledge (Wolter & Yamashita, 2017; Yamashita, 2018). This indicates that the participants did not observe the available L2 nodes which determined the actual sense of

the L2 collocates they translated from the L1 words. Furthermore, it is possible that the participants may not have known the actual meaning of the most typical L2 collocates as constrained by their collocational co-texts.

Second, processing incongruent collocations accurately is complex and cognitively demanding (Wolter & Gyllstad, 2013; Yamashita & Jiang, 2010). Much evidence indicates that incongruent collocations are difficult to acquire and process, even for advanced L2 learners. While EFL learners can process congruent collocations effortlessly by triggering L1 translation equivalents (Wolter & Yamashita, 2017), this psycholinguistic process is not straightforward for mastering incongruent collocations. When encountering incongruent collocations, EFL learners undergo a complex psychological process (Hoey, 2005). That is, they are predisposed to perform new interlingual connections and restructure their L2 lexical network to arrive at common collocations (Wolter, 2006). This explanation suggests that if learners' existing L2 collocational knowledge is not sufficient, they will likely offset this lack by resorting to their L1 lexical knowledge and hence produce infelicitous collocations. This tendency could occur especially when learners perform controlled tasks which force them to attend to the L1 meaning.

Finally, the participants found incongruent collocations difficult due to the fact that providing accurate L2 collocates requires both paradigmatic and syntagmatic knowledge. Students need both types of knowledge as they make semantic and syntactic connections between the L1 and L2 and between the translated L2 equivalents and the given L2 node. For paradigmatic connections, students have to create semantic relations between L1 and L2 words (Cruse, 2006). Regarding syntagmatic connections, moreover, students have to establish lexical interaction between nearby words (Wolter, 2006). These surrounding words are syntactically associated with and semantically determined by each other (Sinclair, 2008). Thus, if L1 and L2 words are semantically incongruent, and students are unaware of their collocating words, there is a high risk they will derive inaccurate L2 words. The results of the present study regarding L1-L2 incongruency illustrated this situation. The participants provided incongruent L2 collocates which were not accurate nor typical (most typical in parentheses), for example, *find problem (encounter)*, *make research (conduct)*, *make awareness (raise)*, *low performance (poor)*, *delicate analysis (careful)*, *highly linked (strongly)*, *only academic (purely)*, and *very successful (highly)*. The results clearly indicate that the

L2 collocates of these incongruent L2 collocations are directly combined from the meanings of the L1 equivalents.

4.3 Number and quality of academic collocations

Table 3 presents the statistics for the collocate types answered by the participants in answer to the third research question. Notably, of the 569 types, the ME group generated 351 types, making up 64.69%. This percentage is almost two times higher than that of the types given by the LE group.

Table 3

Comparison of collocate types by academic experience

Participants	Collocability			χ²	p
	Yes (%)	No (%)	Total (%)		
LE	48 (22.02)	170 (77.98)	218 (38.31)	0.09	0.77*
ME	81 (23.08)	270 (76.92)	351 (61.69)		
Total (%)	129 (22.67)	440 (77.33)	569 (100.00)		

*p > .05.

The data in Table 3 show that the ME participants performed slightly better than the LE participants on the number and quality of academic collocations. That is, they tended to prefer the most typical collocations. However, the chi-square test revealed that there was no statistically significant difference in the number and quality of academic collocations favored by the LE and ME groups (χ^2 = 0.09, *df* = 1, *p* > .05). It can be seen that the number of supplied collocates was positively correlated with the number of incorrect answers. This is evidenced in the case of the ME group who attempted to demonstrate their L2 lexical knowledge by providing multiple L2 collocates, yet their erroneous choices increased according to the number of responses. This suggests that a good knowledge of individual lexis does not adequately contribute to collocation acquisition.

The participants from both groups provided 569 L2 collocate types, 440 (77.33%) of which were of poor quality and hence not acceptable. Among these non-acceptable collocates, most of them carried equivalent L1 meanings, yet they were not typical combinations. For example, they supplied many unusual collocates that shared the L1 sense of *tòk-tàm*, such

as *low*, *decreased*, *declined*, *dropped*, *plummeting*, and *lower-quality*. Only three ME participants gave the acceptable collocate *poor* to modify the node *performance*. As can be seen, the participants attempted to provide stylistically marked collocates, but obviously they were not acceptable. In fact, there were many other questions that received similar answers. Furthermore, we found that the majority of the participants preferred general lexis, such as *very* (126 tokens), *have* (104 tokens), *get* (99 tokens), *make* (79 tokens), *give* (70 tokens), and *do* (58 tokens). These hugely overused responses constituted a large proportion of the infelicitous collocations. This result was also observed among native-English academic writers (Frankenberg-Garcia, 2018). In academic discourse, it is generally acknowledged that general delexical items are not typical. They are indeed high-frequency words, but not strongly associated with the nodes targeted in the present study.

The study revealed that the participants, especially the ME group, produced multiple equivalent L2 collocates for one single L1 word. This seems to suggest that they had developed a repertoire of L2 lexical knowledge to some extent. Despite knowing many L2 equivalents, the results clearly indicate that they had inadequate knowledge of academic L2 collocations. This inadequacy may be attributed to the participants' L1 influence and limited L2 lexicon. First, the participants may have relied too heavily on their L1 lexical and conceptual knowledge which they literally transferred to the L2 collocations (Wolter & Gyllstad, 2013; Yamashita & Jiang, 2010). The participants of both groups may not have realized that the meanings of the collocates they provided were predicated on the L2 rather than the L1 and largely determined by the nodes. Consequently, the participants ended up supplying several L2 collocates which were not acceptable. Second, it is probable that the participants may not have entrenched acceptable L2 collocates in their L2 lexicon. That is, they may not have been (adequately) exposed to them in their learning discourse. Given this situation, they resorted to general delexical items. The choices of delexical items may be due to the fact that they have high-frequency equivalents in most languages (Altenberg & Granger, 2001), probably including Thai. Delexical items are also well-known for carrying little or no meaning (Stubbs, 2001), and they can occur freely with a large number of words. In sum, when forming unfamiliar L2 collocations, EFL learners are likely to draw on delexical items to compensate for typical L2 collocates that they do not know, or are not aware of.

5. Conclusion and Implications

This study explored Thai EFL learners' knowledge of academic collocations focusing on three variables: academic experience, L1-L2 congruency, and collocation quality. The results showed that the ME participants significantly outperformed their LE counterparts on their knowledge of academic collocations. However, we found that both LE and ME participants had greater difficulty with incongruent than congruent collocations, given the inequivalent L1-L2 meanings of the collocates tested in the study. This L1-L2 incongruency predisposed the participants to derive unusual L2 collocates from their L1 lexical knowledge. Of these poor-quality collocates, many were also delexical items supplied by the participants to compensate for their lack of knowledge and awareness of typical L2 collocations.

These results suggest some pedagogical implications for teaching academic collocations. The fact that the students produced a large number of lexis, more than half of which were not acceptable, points to a need for higher-quality vocabulary instruction. It is evident from this study that it is not sufficient for learners to develop their L2 lexicon implicitly; they should be made aware of context and co-text in vocabulary instruction. Thus, teachers should concentrate on teaching lexical items in relation to their respective register and collocational pattern. In practice, students should be assigned to (1) read academic texts that contain typical academic collocations; (2) identify common collocates and syntactic patterns; and (3) observe co-occurrence frequency and collocation typicality by using free-access academic corpora (e.g. BNC, COCA) (https://www.english-corpora.org/coca/). These teaching methods will help learners to realize that when words are used together, they keep certain company, which conveys their actual meaning. This linguistic behavior is typical of academic discourse which is recognized to be lexically dense and highly phrasal.

The results clearly indicate that incongruent L2 collocations are problematic for EFL learners, even for those who have experienced greater exposure. It is therefore more effective for teachers to teach incongruent collocations through explicit methods. Those direct teaching methods should assist students to firmly master incongruent collocations and to process them independently of their L1 lexicon and concepts (Yamashita & Jiang, 2010). To achieve these goals, teachers should raise students' awareness of different L1-L2 forms and meanings (Nesselhauf, 2003) and

provide learning activities with semantic elaborations (Snoder & Reynolds, 2019). In practice, students can learn to translate L2 collocations into L1 equivalents and vice versa, and then consider the meanings of the component words in separation and combination. During this process, they may find out that many L2 collocations are semantically idiosyncratic and inequivalent to their component words and that the true senses of lexical items need to be determined in association with their collocates in context, such as *strongly linked*, *purely academic*, and *poor performance*. In addition, given that many L2 lexical items are polysemous, students should learn multiple collocates for one single node (e.g. Webb & Kagimoto, 2011), such as *poor family* vs. *poor performance.* From these two examples, the node *poor* has two different meanings as determined by its collocates *family* and *performance*. This method is particularly useful for students because they can disambiguate polysemous words and elaborate their specific meanings with respect to their typical collocational behaviors.

However, it may not be sufficient for EFL teachers and students to observe the semantic and syntactic behaviors of L2 collocations. Given the fact that language use is register-specific, it is important that teachers should devote special attention to the in-context functions of academic L2 collocations. There are many academic collocations that co-function to express specific purposes. For example, *significantly different* is used to report statistically meaningful results, *strongly suggest* to interpret research results with heightened confidence, *widely accept* (often in a passive *it*-clause) to indicate commonly shared knowledge, *hardly possible* (often in an *it*-clause) to express modality, *strongly linked* to define a positive association, *mainly focus* to express a specific purpose, *play role* to show a causal relationship, and *strong evidence* to present a convincing proof. By focusing on the functions of L2 collocations in their natural context, students can develop their L2 collocational network to the extent that it is overtly divergent from their L1 lexical network (Wolter, 2006). Subsequently, they can process their acquired L2 collocational knowledge effortlessly in receptive and productive tasks.

6. Limitations and Recommendations

This study has some limitations that need to be addressed and merit further investigation to obtain more comprehensive results. First, the study used only the number of university years as the variable to divide

the participants into two groups to compare their collocational knowledge in relation to their exposure to academic discourse. However, academic experience alone may not fully explain students' knowledge of academic L2 collocations. To complement this variable, therefore, future researchers should examine students' semantic knowledge of general academic words by employing a translation task. Among these words, they should also include component words which will appear in a collocation test. This semantic knowledge can offer valuable insight into students' academic L2 collocation, especially incongruent ones with respect to their proficiency apart from their L1 and learning experience. Translations of collocations and their component words can be compared to obtain more insightful data.

Second, the findings of the present study represent the outcome of a gap-filling translation test. This test required the participants to provide L2 collocates as their responses. While the test allowed us to obtain the participants' answers as the products for examining their collocations, the processes during which they thought and arrived at their answers were not explored. In the present study, the processes of solving the collocation problems encompassed translating the target lexical items from the L1 to L2 and choosing the most typical translation as the L2 collocate. Exploring these processes would be valuable in providing more insight into participants' cognitive mechanisms that trigger the choices of L2 collocates. Therefore, it is suggested that further similar studies employ a retrospective report as a complementary qualitative approach in which participants verbalize their thinking and reasoning processes immediately after they complete a given task (McKay, 2009). This follow-up instrument will corroborate product-based results and hence enable us to delve deeper into their collocational knowledge and how they process congruent and incongruent collocations in relation to L1 equivalents.

Acknowledgements

About the Authors

Waraphorn Khantiwong: A fourth-year undergraduate student in English and Communication Program, Faculty of Liberal Arts, Ubon Ratchathani University, Thailand. Her research interest includes formulaic language in academic discourse.

Atikhom Thienthong: A lecturer in English and Communication Program, Faculty of Liberal Arts, Ubon Ratchathani University, Thailand. His research interests in academic writing are collocation, synonymy, pattern grammar, written discourse, and corpora.

References

Adamska-Sałaciak, A. (2013). Equivalence, synonymy, and sameness of meaning in a bilingual dictionary. *International Journal of Lexicography 26*(3), 329–345. https://doi.org/10.1093/ijl/ect016

Altenberg, B., & Granger, S. (2001). The grammatical and lexical patterning of MAKE in native and non-native student writing. *Applied Linguistics, 22*(2), 173–195. https://doi.org/10.1093/applin/22.2.173

Biber, D., & Gray, B. (2010). *Grammatical complexity in academic English: Linguistic change in writing.* Cambridge University Press.

Chansopha, N. (2018). The use of international business management collocations by Thai EFL speakers. *LEARN Journal: Language Education and Acquisition Research Network, 11*(1), 87–109.

Conklin, K., & Schmitt, N. (2012). The processing of formulaic language. *Annual Review of Applied Linguistics*, *32*, 45–61. https://doi.org/10.1017/S0267190512000074

Cruse, D. A. (2006). *A glossary of semantics and pragmatics*. Edinburgh University Press.

Durrant, P., & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research, 26*(2), 163–188. https://doi.org/10.1177/0267658309349431

Fernández, B. G., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have?: The effects of frequency and amount of exposure. *International Journal of Applied Linguistics, 166*(1), 94–126. doi:10.1075/itl.166.1.03fer

Flowerdew, L. (2012). *Corpora and language education*. Palgrave Macmillan.

Frankenberg-Garcia, A. (2018). Investigating the collocations available to EAP writers. *Journal of English for Academic Purposes*, *35*, 93–104. https://doi.org/10.1016/j.jeap.2018.07.003

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics, 35*(3), 305–627. https://doi.org/10.1093/applin/amt015

Hoey, M. (2005). *Lexical priming: A new theory of words and language.* Routledge.

Hyland, K. (2013). Writing in the university: Education, knowledge and reputation. *Language Teaching, 46*(1), 53–70. https://doi.org/10.1017/S0261444811000036

Huddleston, R., & Pullum, G. K. (2017). *The Cambridge grammar of the English language*. Cambridge University Press.

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners. *English Language Learning, 61*, 647–672. https://doi.org/10.1111/j.1467-9922.2010.00621.x

Lei, L., & Liu, D. (2018). The academic English collocation list: A corpus-driven study. *International Journal of Corpus Linguistics*, *23*(2), 216–243. https://doi.org/10.1075/ijcl.16135.lei

Men, H. (2018). *Vocabulary increase and collocation learning: A corpus-based cross-sectional study of Chinese learners of English.* Springer. doi:10.1007/978-981-10-5822-6

McKay, S. L. (2009). Introspection techniques. In J. Heigham & R. A. Croker (Eds.), *Qualitative research in applied linguistics: A practical introduction* (pp. 220–241). Palgrave Macmillan.

Nagy T., I., Rácz, A., & Vincze, V. (2020). Detecting light verb constructions across languages. *Natural Language Engineering, 26*, 319–348. doi:10.1017/S1351324919000330

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics*, *24*(2), 223–242. https://doi.org/10.1093/applin/24.2.223

Nesselhauf, N. (2005). *Collocations in a learner corpus.* John Benjamins.

Nguyen, T. M. H., & Webb, S. (2016). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research, 21*(3), 1–23. https://doi.org/10.1177/1362168816639619

Peters, E. (2015). The learning burden of collocations: The role of interlexical and intralexical factors. *Language Teaching Research, 20*(1), 1–26. https://doi.org/10.1177/1362168814568131

Phoocharoensil, S. (2013). Cross-linguistic influence: Its impact on L2 English collocation production. *English Language Teaching, 6*(1), 1–10. http://dx.doi.org/10.5539/elt.v6n1p1

Rychlý, P. (2008). *A lexicographer-friendly association score*. https://www.sketchengine.eu/bibliography-of-sketch-engine/

Sinclair, J. (2008). The phrase, the whole phrase, and nothing but the phrase. In S. Granger & F. Meunier (Eds.), *Phraseology: An interdisciplinary perspective* (pp. 407–410). John Benjamins.

Sanguannam, S. (2017). A study on "delexical verb+noun" collocation errors of Thai EFL intermediate and advanced learners. *Journal of Liberal Arts, 17*(1), 59–84.

Siyanova-Chanturia, A. (2015). Collocation in beginner learner writing: A longitudinal study. *System, 53*, 148–160. https://doi.org/10.1016/j.system.2015.07.003

Siyanova-Chanturia, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review, 64*, 429–458. doi:10.3138/cmlr.64.3.429

Siyanova-Chanturia, A., & Spina, S. (2015). Investigation of native speaker and second language learner intuition of collocation frequency. *Language Learning, 65*(3), 533–562. https://doi.org/10.1111/lang.12125

Snoder, P., & Reynolds, B. L. (2019). How dictogloss can facilitate collocation learning in ELT. *ELT Journal*, *73*(1), 41–50. https://doi.org/10.1093/elt/ccy024

Stubbs, M. (2001). *Words and phrases: Corpus studies of lexical semantics.* Blackwell.

Stubbs, M. (2009). The search for units of meaning: Sinclair on empirical semantics. *Applied Linguistics*, *30*(1), 115–137.

Tsai, M-H. (2018). The effects of explicit instruction on L2 learners' acquisition of verb-noun collocations. *Language Teaching Research, 24*(4), 1–25. https://doi.org/10.1177/1362168818795188

Vilkaitė, L., & Schmitt, N. (2019). Reading collocations in an L2: Do collocation processing benefits extend to non-adjacent collocations? *Applied Linguistics, 40*(2), 329–354. https://doi.org/10.1093/applin/amx030

Wang, Y. (2016). *The idiom principle and L1 influence: A contrastive learner-corpus study of delexical verb + noun collocations.* John Benjamins.

Webb, S., & Kagimoto, E. (2011). Learning collocations: Do the number of collocates, position of the node word, and synonymy affect learning? *Applied Linguistics, 23*(3), 259–276. https://doi.org/10.1093/applin/amq051

Wolter, B. (2006). Lexical network structures and L2 vocabulary acquisition: The role of L1 lexical/conceptual knowledge. *Applied Linguistics, 27*(4), 741–747. https://doi.org/10.1093/applin/aml036

Wolter, B., & Gyllstad, H. (2013). Frequency of input and L2 collocational processing: A comparison of congruent and incongruent collocations. *Studies in Second Language Acquisition, 35*(3), 451–482. https://doi.org/10.1017/S0272263113000107

Wolter, B., & Yamashita, J. (2017). Word frequency, collocational frequency, L1 congruency, and proficiency in L2 collocational processing: What accounts for L2 performance? *Studies in Second Language Acquisition, 40*(2), 395–416. doi:10.1017/S0272263117000237

Wray, A. (2002). *Formulaic language and the lexicon.* Cambridge University Press.

Yamashita, J. (2018). Possibility of semantic involvement in the L1-L2 congruency effect in the processing of L2 collocations. *Journal of Second Language Studies, 1*(1), 60–78. https://doi.org/10.1075/jsls.17024.yam

Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL Quarterly, 44*(4), 647–668. https://doi.org/10.5054/tq.2010.235998

Yoon, H. J. (2016). Association strength of verb-noun combinations in experienced NS and less experienced NNS writing: Longitudinal and cross-sectional findings. *Journal of Second Language Writing, 34*, 42–57. https://doi.org/10.1016/j.jslw.2016.11.001

Appendix 1

Frequency and typicality (logDice) of congruent collocations

Collocation	Frequency/logDice	Translation
verb + noun		
provide benefit	73,288/7.32	**hâj** praʔ-jòːt
acquire knowledge	37,803/9.36	**dâj ráp** khwaːm-rúː
receive attention	55,096/8.13	**dâj ráp** khwaːm-sǒn-caj
create opportunity	135,767/8.68	**sâːŋ** ʔoː-kàːt
adjective + noun		
main reason	87,130/9.3	hèːt-phǒn **làk**
primary purpose	28,636/8.46	cùt-praʔ-sǒŋ **sǎm-khan**
great impact	48,113/7.19	phǒn-kraʔ-thóp **jàːŋ mâːk**
necessary skill	21,959/7.63	thák-sàʔ thîː **cam-pen**
adverb + adjective		
vitally important	23,809/9.3	sǎm-khan **jàːŋ-jîŋ**
markedly different	5,537/7.48	tɛːk-tàːŋ jàːŋ **hěn dâj chát**
significantly different	25,732/8.77	tɛːk-tàːŋ **jàːŋ miː naj-jáʔ-sǎm-khan**
hardly possible	1,805/5.67	pen paj dâj **jâːk-maːk**
adverb + verb		
change **rapidly**	48,161/9.91	plìan-plɛːŋ **jàːŋ rûat-rew**
widely accept	25,944/9.2	jɔːm ráp **jàːŋ kwâːŋ-khwǎːŋ**
mainly focus	36,540/9.87	mûŋ nén **pen làk**

Note: The words marked in **bold** are collocates.

Frequency and typicality (logDice) of incongruent collocations

Collocation	Frequency/logDice	Translation
verb + noun		
play role	898,734/12.41	**miː** bòt-bàːt
encounter problem	24,192/7.84	**phóp** pan-hǎː
conduct research	211,884/10.98	**tham** wíʔ-caj
raise awareness	301,139/11.61	**sâːŋ** khwaːm-traʔ-nàk
adjective + noun		
strong evidence	21,562/7.69	làk-thǎːn **thîː chát-ceːn**
poor performance	24,383/8.04	phǒn-kaːn-rian **tòk-tàm**
careful analysis	7,657/6.54	kaːn-wíʔ-khrɔ́ʔ **jàːŋ láʔ-ʔìat**
serious problem	80,184/9.14	pan-hǎː **thîː ráːj-reːŋ**
adverb + adjective		
completely different	88,765/10.18	tɛ̀k-tàːŋ **jàːŋ sîn-chəːŋ**
strongly linked	4,710/7.27	sǎm-phan kan **jàːŋ mâːk**
purely academic	1,405/8.18	dâːn wíʔ-chaː-kaːn **thâw-nán**
highly successful	4,3601/9.23	praʔ-sòp khwaːm-sǎm-rèt **jàːŋ mâːk**
adverb + verb		
fully understand	65,734/9.6	khâw-caj **jàːŋ thòŋ-théː**
strongly suggest	20,882/9.43	chíː hâj hěn **jàːŋ chát-ceːn**
look **closely**	31,000/8.19	sùk-sǎː **jàːŋ láʔ-ʔìat**

Note: The words marked in **bold** are collocates.