

Collocations and local textual functions of quantifiers in learner English essays*

Raksangob Wijitsopon
(Chulalongkorn University)

Wijitsopon, Raksangob. Collocations and local textual functions of quantifiers in learner English essays. *Linguistic Research* 34(1), 1-49. Over decades, corpus linguistic research has yielded a number of theoretical insights into linguistic mechanism of the English language. One of these concepts is “local textual function”, which features a dialectical relationship between lexical items and texts. The concept has been applied to various text types but never to learner writing, even though the production of texts is an important communicative skill emphasized in English language teaching and learning. At the same time, learner corpus research has tended to pay particular attention to lexicogrammatical patterns without relating them to textual output produced by language learners. The present study therefore seeks to fill in these gaps through an application of the text-lexicogrammar theoretical construct to an analysis of a linguistic category that has hardly been dealt with in learner corpus research, i.e. quantifiers. The words “some”, “many” and “every” are among the top 25 function keywords in Thai undergraduates’ English argumentative writing, when compared with a corpus of their native speaker counterparts’. They were hence investigated in terms of their phraseological patterns and functional contributions to the discourse of Thai learner essays. Findings from the analysis shed light on the textlinguistic dimension of the quantifiers in native and non-native speaker learner usage and have pedagogical implications. (Chulalongkorn University)

Keywords local textual functions, collocations, quantifiers, learner corpora, corpus-driven

1. Introduction

It cannot be denied that corpus linguistic research over the past decades has

* The present paper is part of a corpus-driven research project on stylistic differences between Thai and native speaker learners’ English argumentative writing, funded by the Thailand Research Fund (TRF), Commission on Higher Education (CHE) and Chulalongkorn University via the TRF-CHE Research Grant for New Scholar and the Ratchadaphiseksomphot Endowment Fund of Chulalongkorn University (TRG5680084). I would like to express my gratitude to my sponsors for their continued support. I am also grateful to two anonymous reviewers for their suggestions and comments on the earlier version of this paper.

yielded important theoretical insights into the way the English language works. That is, meaning in language is created through co-occurrence patterns among lexical items and lexis and grammar cannot be treated separately as they both together form units of meaning. These concepts have been applied extensively to various areas in applied linguistics, including studies on learner English. A great deal of learner corpus research has centered around lexicogrammatical patterns in native and non-native learner English writing (e.g., Lorenz 1999; Nessellhauf 2005; Chen and Baker 2010; Li and Schmitt 2010; Wei and Lei 2011). Granger (2015) observes that these corpus-based concepts have enriched language acquisition and interlanguage research, which has long focused on morphological and syntactic domains, with findings and discoveries on phraseological phenomena in learner English.

However, recent corpus linguistic studies argue that the form-meaning relationship of linguistic features goes beyond co-occurrence between lexicogrammatical items and encompasses textual features that contextualize those phrasal patterns (Stubbs 2015; Hoey 2005). As Mahlberg (2009: 267) nicely comments:

Having made clear that grammar and lexis cannot be separated, corpus research needs to go further and pay appropriate attention to the textlinguistic dimension of meaning. Then a corpus is truly exploited as a collection of texts and not only as a source of concordance lines.

The present study adopts this theoretical stance, through the concept of “local textual function” (Mahlberg 2005, 2009, 2013), and applies it to an analysis of three quantifiers “some”, “many” and “every”, which have been found to be keywords in Thai undergraduates’ essays when compared with their native speaker counterparts (see Section 4.2). This is in order to demonstrate how the recently developed theoretical concept can benefit an applied linguistic study of learner English and, at the same time, to expand the scope of learner corpus research, which has over decades centered on phraseological patterns including collocation and lexical bundles (Paquot and Granger 2012), to that which incorporates textual dimension of lexical items in the analysis of learner data. Moreover, quantifiers have rarely been discussed in relation to its textlinguistic dimension in non-native learner writing even though they seem to have received substantial attention in second language

acquisition and interlanguage research, as exemplified in such studies as Katsos *et al.* (2016), Chu, Gabriele, and Minai (2014) and Marsden (2009). The corpus-driven stance adopted in this study thereby suggests that some quantifiers, with their density and variety in Thai EFL students' writing, deserve to be examined in discourse-functional terms.

To these ends, the theoretical framework "local textual function" is applied and will be spelled out first. Then, an overview of quantifiers is provided, followed by descriptions of corpus-driven methods and descriptive tools which have led to a focus on the quantifiers "some", "many" and "every" and which have been employed for a qualitative analysis of the given words. The three keywords are then analysed in detail and discussed before a conclusion that remarks on both pedagogical implications and theoretical contributions of the present study is given.

2. Theoretical framework: From collocation to local textual function

Corpus linguistics is often viewed as "nothing but methodology" (McEnery and Wilson 2001: 1) because it features the use of a corpus as a means of examination of various aspects of language, e.g., lexis, grammar, discourse, but does not define a particular aspect of language that requires description in the way that syntax and semantics focus on grammatical structure and meanings in language, respectively (McEnery and Wilson 2001: 2). However, work in corpus linguistics over the past decades has shown that the use of corpora can offer a distinctive way of looking at language, which allows a description of language from new points of views. Specifically, as Tognini-Bonelli (2001: 2) argues, a corpus linguistic framework involves a description of language that privileges lexis, its observable patterns and their relations to meanings or functions. Well-known corpus-driven descriptive categories include different kinds of co-occurrence patterns among lexical items, such as collocation (co-occurrence between lexical items), colligation (co-occurrence between a lexical item and grammatical categories), semantic preference (co-occurrence between a lexical item and semantic fields), cluster (uninterrupted string of lexical items, also called n-gram, where "n" stands for the number of lexical items in a cluster, say "3-grams" or "5-grams"). Some of these concepts were

adopted in the present study to describe patterns of the quantifiers under study.

However, it must be stressed that various kinds of co-occurrence patterns exemplified above are embedded in the texts that make up a corpus. They are therefore inextricably linked to the texts in question. On the one hand, words and their patterns contribute to textual meanings and organization. For example, they characterize a particular text type, e.g., academic texts vs. conversation (Conrad and Biber 2005), or signal a rhetorical structure, such as the problem-solution pattern (Flowerdew 2008). At the same time, functions that a lexical item performs in a text can be regarded as one of its semantic and pragmatic properties (Hoey 2005; Mahlberg 2009; Hoey and O'Donnell 2015).

To systematically account for the above dialectical relationship between text and lexis, there have been attempts in corpus linguistics to develop a framework that can be applied to explain the relations. One of the concepts proposed is “Local Textual Function” (henceforth LTF) (Mahlberg 2005). LTFs refer to functions of a word which connect it with its context and contribute to the creation of a text. An important characteristic of LTF is that they do not aim to be applicable to any text or type of text. As the name suggests, the functions are “local”, i.e. particular to a specific (set of) lexical items(s) and/or a specific (set of) text(s). Importantly, the description of LTF often involves *ad hoc* categories, since the functional relationship between text and lexical items tends to vary according to textual and contextual factors, including word positions, text genres, domains or contents. The concept of LTF is thus flexible enough to be adopted to investigate various groups of words and texts, ranging from general nouns in a general corpus (Mahlberg 2003), in newspapers (Mahlberg 2009) to clusters in classic novels (Mahlberg 2013). Although LTF of a particular word can vary across different textual varieties and factors, what is important, as Mahlberg (2009) stresses, is that labels of textual functional categories can be derived systematically from the parameters set for the interpretation, e.g., frequency, text genres and sections in a text. To illustrate, the body-part cluster “his head on one side” in Charles Dickens’ novels can serve the *contextualizing* function when it accompanies the description of a situation or activity and can also perform the *highlighting* function when it occurs repeatedly in association with a particular character, highlighting his or her habits or behavior (Mahlberg 2007).

The concept of LTF has never been applied to learner corpora. In fact, as

mentioned earlier, the text-lexicogrammar perspective has not quite been touched upon in learner corpus analysis. The dominant line of enquiry tends to involve the interpersonal dimension of writing, including identity construction, stance expression and writer-reader interaction (e.g., Hyland 2003). Other aspects in textual systemics, e.g., textual properties or thematic structure, have received much less attention and hence should be pursued to advance the field (Flowerdew 2002). The present study therefore applies LTF as a major descriptive tool to investigate the ways in which three quantifiers are used in the creation and organization of English argumentative essays written by Thai undergraduates and their native-speaker counterparts.

3. Quantifiers

Quantifiers have received attention in applied linguistic studies, especially in language acquisition research. Much of the literature in the field concentrates on L2 learners' acquisition of different aspects of quantifiers, particularly quantifier scope (Chu, Gabriele, and Minai 2014; Ionin, Luchkina, and Stoops 2014) and count/mass noun distinction perceived by non-native speaker learners (Lima 2014). They have also been relatively extensively in general linguistic description and English grammar references. They are labeled as a type of determiners but specify nouns in terms of quantity. Quantifiers are usually approached in grammatical or semantic terms; for instance, they are often described with respect to their being used to quantify count or uncount nouns. They also tend to be categorized according to their semantic properties as, for example, assertive or non-assertive determiners (Quirk *et al.* 1985: 383). Biber *et al.* (1999) divide quantifiers into four groups:

- (1) Inclusive, which specifies the whole of a group or mass, namely “all”, “both”, “each” and “every”;
- (2) Large quantity, which specifies a large quantity of noun, namely “many”, “much”, “a lot of”, “lots of”, “a great deal/ number of” and “plenty”;
- (3) Moderate or small quantity, with “some” and “several”, denoting a moderate quantity and “a few”, “few”, “a little” and “little”, indicating

a small quantity

- (4) Arbitrary/ negative member or amount, with “any” and “either”, referring to an arbitrary member of a group or amount of a mass, while “no” and “neither” serving as negative quantifiers for the arbitrary group.

Based on the above categorization, it can be seen that the three keyword quantifiers in Thai students’ essay corpus to be analyzed functionally, “every”, “many” and “some”, belong to groups (1), (2) and (3), respectively. It can thus be interpreted that overall Thai students’ argumentative essays tend to highlight a reference to quantity, when compared with native speaker learners’ essays. The collocation and LTF analyses in Section 5 below will provide a clearer and more precise description of their functional contributions to the essays.

Apart from these grammatical and semantic approaches to quantifiers, Biber *et al.* (1999) touches on the text-lexis relationship by looking at the distribution of quantifiers in different text types. It is shown that there is little variability in their distributions across registers, which suggests that quantifiers are common words in almost any text types. However, there are wide differences in the distribution of individual quantifiers. For example, the words “some” and “many” are the most common quantifiers in academic prose, as they fulfill “the need for expressing guarded generalisations” (Biber *et al.* 1999: 277). An attention to the correlation between lexical distribution and register reflects a quantitative approach to a text-lexis account of quantifiers. However, for the purpose of the present study, it cannot be applied to explain why some quantifiers were used significantly by a group of EFL learners like Thai undergraduates’ because the description relies on a general corpus. Applying such a general usage profile to learner corpora can be problematic since it could possibly lead to an impression that Thai students’ essays sound more academic than native speakers’ as “many” and “some” occur significantly in Thai undergraduates’ writing (see Section 4.2). A more specific kind of framework like LTF, which allows a functional description specific to a particular text type, is thereby needed.

4. Methodology

In this section, information about the corpus data used in the study is provided in 4.1. This is followed by details about descriptive tools and analytical procedures adopted for the quantitative and qualitative analysis of the Thai learner corpus.

4.1 Data

Two corpora are mainly involved in the present study: (1) a corpus of Thai undergraduates' English argumentative essays (henceforth THAI) and (2) a corpus of native-speaker learners' argumentative essays (henceforth NATIVE). NATIVE contains the university student essay component of the Louvain Corpus of Native English Essays (LOCNESS), comprising written output from both British and American students, with 264,095 tokens, while THAI holds 100 argumentative essays with 41,821 tokens. The size difference between the two corpora fits in Barber Sardinha's (2004 cited in Scott and Tribble 2006) suggestion that a reference corpus should be at least five times larger than the text under investigation. A native-speaker learner corpus was used as a reference corpus to extract keywords in THAI because it is the closest possible to THAI in terms of text type and text producers and because nativeness seems appropriate as a benchmark for an investigation of non-native students' performance (Paquot 2010; Mukherjee 2005). The essays in THAI were taken from five universities in Thailand, each of which contributed its top 20 best essays written by their undergraduate students as part of their final exams in English composition courses. The rubrics used in the marking process involve organization, contents, cohesion and coherence, grammar and word choices. The reason why final exam essays were chosen to make up the corpus is that it was written after the students had learnt and been trained to write argumentative essays for a semester of 15 weeks. These essays were graded by lecturers in composition courses at each university. The top 20 highest-scored essays from each university were recruited in order to maximize the quality of the writing by Thai EFL students to be compared with native speaker learners, so that issues of errors are minimized and stylistic features can be reflected relatively clearly. The topics of the essays in THAI and NATIVE are mostly different, with such topics as single-sex school, animal use for medical experiment, and privacy issues in THAI

and fox-hunting, boxing, capital punishment, abortion and politics in NATIVE. The topic that both corpora have in common is computer and technology and both contain literary criticism essays. The difference in essay topics is an important factor to be considered in most of the studies using learner corpora, as observed by Granger and Paquot (2010), and hence will be taken into account when interpreting results of the study.

4.2 Keyword

The term “keyword” has been used in various senses. In this study, I follow Scott and Tribble’s (2006: 55) view of “keywords” as a textual concept: “lexical items of significance to a text in question, whether as “aboutness” indicators or as stylistic indicators, because of their “unusual frequency in comparison with a reference corpus of some suitable kind”. Based on this definition, it can be said that a comparison between corpora and word frequency are essential to an identification of keywords of a text. To extract key words, corpus software computes: (1) frequencies of word forms in a text under investigation, (2) the number of running words in the text, (3) frequencies of word forms in a reference corpus and (4) the number of running words in the reference corpus. All these values are then cross-tabulated. Analysts need to set the parameters for words to be identified by the software as key. These parameters are: (1) the minimum frequency of words to appear in the list as key words (2) statistical tests of significance and (3) the probability (i.e. p) value, which suggests the probability that a word is “key” due to chance alone. The keyness of words is thus measured statistically. The corpus software Wmatrix (Rayson 2003), which was used for the present study, offers log-likelihood as a statistical measure to extract keywords. As Leech *et al.* (2001) and Rayson *et al.* (2004) point out, the log-likelihood can deal with a comparison between small data and hence seems appropriate for an analysis of the two learner corpora described above, whose sizes are considered relatively small in corpus linguistic research. The minimum frequency set for an extraction of keywords in the present study is 50 and the log-likelihood critical value is 15.13, amounting to the p value of $p \leq 0.0001$. This is because THAI and NATIVE, though similar in genre, are largely different in terms of their contents and hence a relatively high cut-off point can extract words that are particularly key in THAI. Also, as Rayson *et al.*

(2004) suggest, an optimal log-likelihood value for corpus linguistic work is the critical value of 15.13.

According to Scott and Tribble (2006: 55), three kinds of words usually come out of a comparison as keywords: (1) proper nouns, (2) words that “human beings would recognise” as key, which tend to indicate a text’s “aboutness”, and (3) words that are not usually identified consciously by readers as key but nonetheless occur in significantly high frequencies and so can be indicators of the style of a text, rather than of its content. In other words, a word form which is often repeated within the text under investigation is very likely to be indicative of its meaning and style. Scott and Tribble (2006) also note that the fact that words in category (3), often function words, are identified as key should prompt an analyst to go back to the text and investigate why those words occur with unusual frequencies.

Based on the criteria spelled out above, a total of 82 keywords are shown to mark THAI off from NATIVE (see Appendix 1), 51 of which are content words, most of which suggest subject matters of the essays in THAI (see Appendix 2). The other 31 are function words, which point to the style of THAI learner argumentative writing. These function keywords are shown in Table 1 below.

Table 1. Function keywords in THAI

No.	Word	Frequency in THAI	Per cent	Frequency in NATIVE	Per cent	LL value
1	can	440	1.1	988	0.39	271.2
2	might	111	0.28	70	0.03	218.47
3	some	225	0.56	402	0.16	191.36
4	you	222	0.55	473	0.19	148.44
5	do	261	0.65	662	0.26	130.36
6	they	496	1.23	1700	0.68	122.23
7	your	81	0.2	95	0.04	105.89
8	will	288	0.72	851	0.34	104.76
9	we	273	0.68	795	0.32	102.68
10	have_to	100	0.25	162	0.06	95.49
11	their	349	0.87	1170	0.47	91.74
12	them	171	0.43	479	0.19	70.05
13	or	273	0.68	941	0.38	66.14

10 Raksangob Wijitsopon

14	i	233	0.58	768	0.31	64.24
15	according_to	67	0.17	113	0.05	61.09
16	it	552	1.37	2399	0.96	53.92
17	are	470	1.17	1992	0.8	52.05
18	many	178	0.44	577	0.23	51.37
19	more	232	0.58	831	0.33	49.85
20	about	143	0.36	432	0.17	49.37
21	every	67	0.16	141	0.05	48.47
22	for	507	1.26	2224	0.89	47.16
23	such_as	69	0.17	155	0.06	42.5
24	so	123	0.31	417	0.17	31.3
25	than	96	0.24	305	0.12	29.22
26	have	303	0.75	1363	0.54	24.52
27	should	150	0.37	581	0.23	24.46
28	our	119	0.3	470	0.19	17.97
29	us	58	0.14	185	0.07	17.47
30	my	61	0.15	200	0.08	17.07
31	because	139	0.35	589	0.24	15.41

The keywords in the above table can be put into six groups, according to their parts of speech:

- (5) modal and semi-modal verbs, i.e. “can”, “might”, “will”, “have to” and “should”
- (6) auxiliary verbs, i.e. “do”, “are” and “have”
- (7) pronouns, i.e. “you”, “they”, “your”, “we”, “their”, “them”, “our”, “us”, “my”, “I” and “it”
- (8) prepositions, i.e. “according to”, “about”, “for”, “such as” and “than”
- (9) conjunctions, i.e. “or”, “so” and “because”

(10) quantifiers, i.e. “some”, “many”, “more” and “every”

The first five categories above have been addressed relatively extensively and shown to mark non-native learner writing style (Hinkel 2002). Quantifiers, on the other hand, have rarely been dealt with in the literature of learner written discourse. It is therefore worth examining in detail the ways in which the three quantifiers, “some”, “many” and “every”, contribute to Thai undergraduates’ textual output. It should be noted here that the quantifier “more” is not studied in the present study because its comparative form seems to obviously signal its role in the creation of the comparative rhetorics in Thai students’ corpus data.

4.3 Collocation and local textual function analysis

To explain text-functional significance of the three quantifiers qualitatively, two major steps were taken. First, two-grams of each quantifier were extracted via AntConc (Anthony 2014). This is in order to filter out those that are used as a pronoun, characterised by the node quantifier followed by a verb or preposition, such as “some might” or “many of”. The length of clusters was set at two words in order to recruit as many repeated phrases as possible from a small corpus like THAI and thereby maximize the number of phrases to go through a detailed investigation of each quantifier. The clusters to come into a list of collocations were required to occur at least twice in an essay. Again, this is in order to maximize the number of noun types that are used repeatedly in the two corpora for data analysis. Concordance lines of each two-gram were then further examined to identify co-textual patterns that occur beyond the span of two-word phrases. This first step is generally considered a conventional corpus linguistic approach to language, wherein patterns of co-occurrences among lexical items are central to a usage description of a lexical item. The collocational patterns of these quantifiers in THAI were then compared with their NATIVE counterparts. Although the present study aims at investigating functional significance of the keyword quantifiers in Thai students’ writing, a comparison with co-occurrence patterns in NATIVE has been found to shed light on some distinctive features of Thai students’ uses of the three quantifiers in their essays, which might have not been spotted without a look at a reference corpus. Two general corpora, the British National Corpus (BNC) and the Corpus of

Contemporary American English (COCA), were also consulted to examine statistically significant collocates of the quantifiers. This is in order to find out to what extent collocational patterns found in the two learner corpora correlate with or deviate from general usage norm of the three quantifiers. To ensure that the general corpora are as comparable to the learner corpora as possible, only the academic component in each of the corpora was selected to compare with THAI and NATIVE. It must be noted that noun collocates extracted from the general corpora are statistically significant collocates, unlike those frequency-based ones from the learner corpora, since the BNC and COCA are much larger in size.

However, as mentioned earlier, these collocational patterns do not stand alone but are part of the texts in the corpora. In order to explain their roles in the essays in THAI, another step was conducted. It involved an analysis of textual environment in which the collocational patterns occur. The major parameters used in the present study include textual positions of the clusters in question and an organizational pattern of the argumentative essay, which is divided by Hyland (1990) into three major stages, as follows:

- (11) The “thesis” stage, which aims to introduce the proposition to be argued. This part often consists of statements that capture the reader’s attention, contextualize the topic, briefly provide support of proposition and signpost the direction of the essay. In Thailand’s EFL writing class, students are often taught to provide a thesis statement at the very end of the introductory paragraph to spell out their positions in an argumentative essay and signal the essay direction.
- (12) The “argument” stage, which aims to discuss grounds for the thesis. This part of the essay mainly involves making claims that explain why the writer accepts and/ or refutes a particular proposition and providing reinforcement relevant to the claims. In Thailand’s EFL writing course, the students have been taught to make claims through a topic sentence that starts each body paragraph. The topic sentence is then followed by statements that provide information in support of the topic sentence.

- (13) The “conclusion” stage, which aims to affirm the validity of the thesis and consolidate the arguments that have been put forward. This is often achieved through restating the proposition in the “thesis” stage, synthesizing the claims in the “argument” stage and closing with remarks that widen the context of the proposition.

Therefore, in this study LTF provides a bottom-up approach to an *ad hoc* functional analysis of the keyword quantifiers, starting from identification of their co-occurrence patterns and textual positions before these textual patterns are related to generic features of the argumentative essay.

It must be noted that while it seems insightful to examine how frequently a quantifier performs a particular local textual function¹, the collocational and LTF analyses of each quantifier in THAI in the present study were conducted for a qualitative purpose, i.e. it mainly aims to describe what textual patterns and functions the quantifiers tend to have in the creation and organization of Thai undergraduates’ essays. This is because THAI is a relatively small corpus of over 40,000 tokens, as it contains only essays written by a group of students, whose writing skills are rated as above average Thai undergraduates (see Section 4.1). Nevertheless, attempts were made to provide quantitative information where appropriate. The figures and percentages, however, should be interpreted cautiously and it is acknowledged that quantitative information is needed in future LTF analysis of lexical items, using a larger corpus.

5. Analysis and results

This section reports on results from a qualitative analysis of the three quantifiers, “some”, “many” and “every”, respectively. The results are divided into two sub-sections in accordance with the analytical steps taken: collocations and local textual functions of each quantifier.

¹ Studies on pragmatics or functional usage of a lexical item with quantitative information can be found, for example, in Bondi (2008).

5.1 Some

The quantifier “some” is the most key of all quantifiers that turn up as keywords in THAI, which means that it is of particular importance in characterizing Thai undergraduates’ English argumentative writing and in differentiating THAI from NATIVE. According to Quirk *et al.* (1985), the determiner “some” indicates reference to a specifiable, though indefinite, quantity or amount of something. It is also noted that “some” is generally used to quantify plural count or uncount nouns but it can also occur with singular count nouns, “particularly temporal nouns, such as “some day”. With other singular nouns, “some” is considered less usual, and “has the meaning ‘a certain’ or ‘some...or other’, which underlines the indefinite or ‘unknown’ quality of the referent” (Quirk *et al.* 1985: 384). It is this colligation pattern of “some”, i.e. its co-occurrence with different grammatical types of nouns, that marks differences in the use of “some” in THAI from its three native speaker corpora. This is demonstrated and discussed in detail below.

5.1.1 Collocations of “some”

The table below presents four lists of noun collocates in the four corpora. The numbers in parentheses show frequencies of their co-occurrence patterns in the learner corpora while the collocates in the general corpora are statistically significant, as explained in Section 3 above. Note that the top 20 most significant collocates of “some” in the BNC and COCA are listed so that a general tendency of collocational patterns of the quantifier can be observed and also compared with those on the NATIVE list.

Table 2. Noun collocates of “some” in THAI, NATIVE, BNC and COCA

No.	THAI	NATIVE	BNC	COCA
1	people (72)	people (50)	cases	students
2	students (10)	cases (20)	extent	cases
3	information (9)	kind (13)	kind	time
4	parents (4)	sort (10)	degree	people
5	schools (3)	way (8)	way	years
6	benefits (2)	Americans (8)	sort	form
7	disadvantages (2)	time (7)	form	studies

8	scientists (2)	extent (5)	respects	degree
9	subjects (2)	ways (5)	areas	extent
10		women (5)	time	way
11		areas (4)	detail	ways
12		form (4)	examples	children
13		loss (4)	aspects	teachers
14		controversies (3)	people	kind
15		critics (3)	evidence	countries
16		states (3)	years	point
17		cities (2)	writers	evidence
18		distance (2)	indication	research
19		reason (2)	countries	states
20		students (2)	ways	areas

Based on the above lists, a couple of patterns can be spotted. First, although “some” occurs significantly in THAI, with a frequency of 225 tokens (see Table 1 above), it is used with a relatively limited range of nouns; only 9 types of noun collocates were found to occur repeatedly in the essays. The most remarkable point is that the collocational pattern of “some” in THAI concentrates on its co-occurrence with the general noun “people”, whose frequency is more than seven times higher than that of the second runner-up “some students” and even greater than the same corresponding pattern, also ranked first, in NATIVE. It should be noted here that the general noun “people” is also a keyword in THAI (ranked 25 on the keyword list shown in Appendix 1). The fact that both keywords, “some” and “people”, constitute a collocational pattern with this top frequency suggests that the cluster plays a particularly important role in Thai students’ essays, the point which I discuss in terms of its local textual functions below.

The other point worth of note deals with differences between the kind of noun collocates on the THAI and NATIVE lists. With two exceptions of “some people” and “some students”, no clusters on either list are similar. On the one hand, this is largely due to differences in topics and contents of the essays in the two corpora. The occurrences of all noun collocates in THAI, except “some people”, can be attributed to the topics of the essays as outlined above in Section 4, e.g., “some students”, “some information” and “some schools”. A number of noun collocates of “some” in NATIVE also signal essay topics in the corpus, e.g., “some women”, “some states” and “some patients” but their frequencies are not in the top five places on the list. Those that occupy the top 5 highest frequencies, but do not occur at all

in THAI, include nouns that are not related to contents of the essays but general abstract nouns that are part of such formulaic expressions as: (in) “some cases”, “some kind” (of), “some sort” (of), (in) “some way(s), and (to) “some extent”. These phrases enable a writer to avoid providing precise information. The extracts from native speaker learners’ essays below illustrate such uses:

- (a) *One argument against the complete legalization of euthanasia is that the people of America are strongly opposed to the killing of human beings (and in some cases, animals).*
- (b) *Once you reach this stage you can consider yourself a habitual offender or someone that needs some kind of jail sentence, punishment or some type of deterrent to stop you if not slow you down.*
- (c) *European cooperation in the field of defence can be both impressive and effective. Whilst I would be reluctant to hand over all control to defence to any central body, I can see a future for some form of European rapid reaction force.*

A glance at the lists of collocates in the general corpora reveals a pattern similar to native speaker learners as the same singular general nouns are statistically significant collocates of the quantifier in the BNC and COCA, including “cases”, “kind”, “form”, “extent” and “sort”. This suggests a tendency in which “some” in native speaker’s discourse is used as part of formulaic expressions of hedging or vagueness. Interestingly, these nouns do not turn up on the THAI list. This points to a marked difference between Thai undergraduates and native speakers in that the former use “some” significantly in their essays in collocation with plural countable nouns to express the plurality of entities, most of which are relevant to the contents of the essays, whereas the native speaker corpora tend to feature the use of “some” as part of a formulaic marker of their stances towards the subjects under discussion, in which case “some” tends to co-occur with singular general nouns. This in a way contradicts with Quirk *et al.*’s (1985: 384) observation cited earlier that “some” can occur with singular count nouns, but “particularly [with] temporal nouns, such as “some day”. With other singular nouns, “some” is considered less usual”. The corpus data employed in the present study suggest that the case is not only far from being “less usual” but actually significantly common.

A look at the collocational patterns alone, however, does not help explain in what ways an indefinite reference to the plurality of entities is essential to Thai students' writing to the extent that "some" becomes one of the top three function keywords. A further analysis of local textual functions of the clusters is therefore needed to reveal the relationship between the lexicogrammatical patterns of "some" and the texts in which they are embedded.

5.1.2 Local textual functions of "some"

Based on an analysis of a larger textual environment than concordance lines, with particular attention to positions in clauses, paragraphs and essays, textual functions of the above collocational patterns of "some" can be spelled out as follows.

5.1.2.1 Introducing counterarguments

The analysis reveals that "some" is mainly used to signal differing views. In many cases, these viewpoints are in contrast with the writer's. Interestingly, this major textual function of "some" in THAI is associated with a particular phraseological pattern. That is, all the 88 two-word clusters of "some" that contain noun collocates denoting people, i.e. "people" (72 cases), "students" (10 cases), "parents" (4 cases) and "scientists" (2 cases), invariably occupy the theme position of a clause (Halliday 1985), whether it is the main or subordinate one. Out of these 88 instances, a total of 54 instances (61.36%) contain rhemes whose main verbs denote verbal or mental acts, e.g., "say", "argue", and "think", which report opinions or arguments of a thematic group of people. These clauses thereby constitute a core phraseological pattern: "some + people noun + speech/ cognition verb. In many cases, this phrasal sequence co-occurs with modal verbs, the conjunction "that" and contrastive markers, e.g., "although" and "however". Following Stubbs' (2007) observation that the English language is largely characterised by recurrent phraseological patterns that contain a fixed core item with a number of variants, the major phraseology of "some" in THAI can be translated into the following pattern, where the capital letters represent the core element and the small letters in parentheses variants of the phraseology:

(contrastive marker +) SOME + PEOPLE NOUN + (modal verb) +
SPEECH/ COGNITION VERB (+ that)

This pattern can be illustrated in the sample concordance lines below:

oid the bad hand writing, such as I. However, fit student more than the opposite. In addition, tion really better than mixed gender education? e-mail. You can easily talk with others people. ongs etc. it helps people to manage their data. n lead to high income and good life. However, ffective to save the whales or not. Even though ng they want. On the other side of this debate, and opportunity of local people. To begin with, her it should be banned or not. However, while nd of prosecution is extremely wrong. Although something is distracted and irritating enough.	some parents argue that because of computers, their some parents claim that they don't see any harms Some parents may think that sending their children o mail. Some people argued that computers make children become Some people argued that computers are not some people believe that elite universities has some people believe that what Sea Shepherds is some people claim that students from single-gender some people claim that factory farming has an some people claim that factory farming produces some scientists claim that they use animal Some students might not be pleased with the restriction
--	--

Figure 1. Concordance lines of the core phraseology of “some”

Upon an examination of a wider context, it is found that these clauses serve to introduce counterarguments against the writer’s views or ideas that contrast with those expressed earlier in the essays. These differences are sometimes even realized explicitly through the co-occurrence between the core pattern and contrastive markers. The following extracts from Thai students’ essays illustrate this function of “some”.

- (14) *However, some parents argue that because of computers, their children are addicted to social networking and games. This opinion can really influence the general view on computers. But recently, there are parental control functions in most of the computers that can limit the user not to access some programs such as games. It can set timet to prevent them from spending too much time on computers.*
- (15) *Some people might say that when students start to work in the business world, they have to work with something harder than the homework. Yes, maybe it's true, but even when I consider it, I still think that teachers*

should reduce the amount of homework because we are students and the most important thing for us is to study, to concentrate on what the teacher says in the classroom.

In the above examples, the student writers' references to "some + people noun" project an existence of a number of people, whose opinions are reported and challenged later in the paragraphs. The word "some" thereby serves to distance the writer from the specified groups of people and their views.

It must be noted that of all 54 instances of the phraseological pattern spelled out above, as many as 50 cases (92.59%) are made up by the phrase "some people". Given that the total frequency of "some people" is 72, the use of "some people" in this phraseology thereby contributes to 69.44% of its all occurrences. It can therefore be said that the remarkably high frequency of "some people" noted earlier in 5.1.1 is largely due to the fact that this collocation serves as a central phrasal device for Thai learners to introduce counterarguments or differing views in their argumentative essays.

Interestingly, this phraseology and its textual function is also found in NATIVE.

(16) *Education may be the best answer to our problem, but some people feel that the death penalty can stop increased crime.*

(17) *Overcrowded prisons is another concern of pro-legalization advocates. Since the latest war on drugs began in the 1980 s, the nation's prison population has more than doubled. The majority of these prisoners are first time offenders who have no history of violence. Some people would argue that this is for purposes of crime control. This function is not being accomplished however. There has been no decrease in the use of marijuana.*

While this points to the same kind of usage between Thai and native speaker learners', it must be noted that data from NATIVE exhibit a wider range of phraseological patterns used for presenting differing views than those found in THAI. The following excerpts from NATIVE illustrate how "some people" is used in various phraseological units to present differing ideas or counterarguments.

- (18) *Respect is not an easy thing to accomplish. Being a talented basketball player and being women does not seem to mix in the eyes of some people. There are still people out there that believe women are not feminine because of their athletic ability. They don't see women as athletes but as sex objects (at least we don't belong in the kitchen anymore).*
- (19) *A weakness in this argument which develops is that some of the reasoning seems a bit steep, like the idea of suicide becoming as popular as fast food. Thus some people will definitely be turned off by this argument.*
- (20) *These so called images, although maybe pleasant to some people, are not always realistic.*

In addition, upon an examination of textual positions of the above core phraseology of “some”, it was found that among all the 54 phrases, 32 of them, or 59.26%, occur at the very beginning of a body paragraph, which is preceded by paragraphs expressing the writer's arguments on the topic. By contrast, the majority of this phrase in NATIVE, 29 out of 40 (72.5%) were found in the middle of a paragraph and only 8 instances (20%) occur at the very beginning of a paragraph. This distributional pattern suggests that the core phraseology “SOME + PEOPLE NOUN + SPEECH/ COGNITION VERB” was often used by Thai undergraduates not only to introduce opposing views but also to express the main idea of a body paragraph and to organize an essay into the argument - counter-argument structure. On the other hand, native speaker learners seem to prefer to use counter-arguments to develop the main point they discuss in a particular paragraph, rather than giving them substantial weight as a separate paragraph in the way Thai students do. The following extract illustrates this disparity between Thai and native speaker learners' uses of this phraseology.

THAI

Some people might think that voyeurism affects only famous people or celebrities. However, in fact, voyeurism affects everyone, both famous and

non-famous, according to the Guardian's article "Creepshots and revenge porn: how paparazzi culture affects women" (Cochrane, par. 1). Since everyone can share everything in the Internet age, not only famous people like the Duchess of Cambridge, Britney Spear, or Emma Watson who were photographed unconsciously and were published, but also any ordinary women from anywhere (Cochrane, par. 5,7).

NATIVE

So the teams with more money are usually the teams with the best players. *An example of this would be Deion Sanders. He originally played for the San Francisco Forty-Niners, but whenever the Dallas Cowboys offered him more money he decided to play with them. With a salary cap these players would not be able to do this. One reason is because each player would make a set amount each year. Some people would think that this was unfair because the most popular players should make more than the players that sit the bench. But in a way this is wrong, giving more money to a player because he is popular would be unfair. Why should Deion Sanders make more money than Andre Risen just because he is more popular? They both are working the same job, so why should Deion make more money?*

Based on the above detailed analysis, it can be seen that the use of "some" in THAI for this particular textual function is relatively formulaic, i.e. it occurs in a specific, almost entirely fixed, phraselogical pattern and textual position. This might be due to an influence of teaching.

It must be noted that apart from being mainly used to introduce counter-arguments, the quantifier was also found to indicate differences in aspects other than opinions, though less commonly, such as those in experiences and (re)actions among individual varieties. Unlike those that introduce contrasting views, the clusters that signal differences in other aspects do not come in the form of a long extended phraseology, which comprises the theme-rheme clausal elements. The extract from THAI below illustrates how a reference to "some people" suggests the way in which "people in general" is treated as a separate group from "terrorists".

While it is true that terrorists might use the email to commit an act of

terrorism, some people use it to store their intimate conversations with their lover as in the case of Paula Broadwell and her affair with formerly CIA director, David H. Petraeus. If this is the case then shouldn't the government respect a person's privacy?

Based on the above textual evidence, it can thus be stated that “some” is significantly used in THAI as a major lexical item of a formulaic phraseological construction that introduces counter-arguments or presents differences between varieties of entities in Thai students’ argumentation. The use of “some” allows the writers to differentiate groups of people or distance themselves from those who have different views.

5.1.2.2 Supporting claims

The functional label “supporting claims” may be considered rather broad and general. However, it was opted for to capture various possible kinds of rhetorical strategies students used to support the main points they made through a reference to “some”. This includes giving reasons, examples for or elaboration of the writers’ observations. The quantifier “some” is used as part of such supporting statements. The collocations of “some” in this case is often found to co-occur with such expressions as “for example” and “because”, as illustrated below.

(21) However, final exam cannot measure anything. This is because some students just simply remember long enough for exam.

(22) Moreover, if they know each other well enough, they will have the protection for themselves. For example, some girls have never been with boys until she graduates so she will not know about ‘boys’ or ‘men’ how nice they are or how dangerous they are.

In fact, it is the textual positions and cohesive ties between statements that mainly point to this claim-supporting function of “some”. That is, while “some” that introduces counterarguments is often found at the very beginning of a body paragraph in THAI, the use of “some” as part of the supporting information tend to

be found in subsequent statements of a body paragraph. In such cases, the sentences with “some” serve to elaborate or clarify the writer’s claims. Moreover, the use of “some” for this purpose is less formulaic than that for introducing counterarguments. This is reflected through the fact that this function is realized through the collocation of “some” with various plural nouns, both animate and inanimate. Though not being precise, the reference to “some” helps give weight to the information in the sense that the quantity of entities involved in the subject is not too small to back up the claim. The excerpts taken from THAI below illustrate this point.

Nowadays, technology has taken parts in education system. Some schools provide classrooms with computers, projector or slide show screen. Some schools encourage students to use laptops and also provide wireless internet for them.

This kind of usage is also found in native speaker learners’ writing, as shown below.

From the use of guns, people are dying at an alarming rate, daily. Therefore, steps for gun control must be taken. There are some segments of society that are not taking gun control seriously, in light of these obstacles, Gun Control needs to implemented because it would mainly reduce crime, save lives and lessen the fears of innocent ones.

There are also only a few other textual patterns and uses of “some” in THAI but they are not dominant enough to be classified into a separate category. A couple of them are used in a thesis statement, the final sentence of the introductory paragraph, denoting an indefinite number of major arguments to be developed in the body paragraphs. For instance:

[...] Although online education brings about equality among all people to gain access to knowledge whenever they want easily, there are some disadvantages that Chulalongkorn University, the most prestigious and accepted university in Thailand, should consider if providing course contents online for free.

It seems then that “some” occurs significantly in THAI because it is mainly used to introduce counter-arguments and to provide supporting information. The fact that it occurs so frequently that it becomes the most key quantifier for two major purposes suggests that its use in THAI is rather limited, both in form and function.

5.2 Many

The quantifier “many” has been studied quite extensively in linguistics, especially in terms of its semantic and syntactic properties (e.g., Huddleston and Pullum 2002; Partee 1989). Its usage profile has only recently been studied in detail by Dichtel (2016), which argues that “many”, when used unmodified in affirmative sentences, denotes not only a large quantity but also diversity of items, as reflected by its frequent collocation with count nouns of settings and abstract concepts. The present study will explore whether this property of “many” is found in learner usage and extend the description of its lexicogrammatical patterns to textual functions in learner argumentative essays.

5.2.1 Collocations of “many”

There are a total of 15 noun collocates adjacent to “many” in THAI. Table 3 below presents the top 20 noun collocates of “many” in the four corpora, with their corresponding frequencies in parentheses.

Table 3. Top 20 two-word clusters of “many” in THAI, NATIVE, BNC and COCA.

No.	THAI	NATIVE	BNC	COCA
1	people (44)	people (88)	cases	people
2	ways (12)	cases (16)	people	students
3	animals (9)	times (15)	ways	years
4	things (9)	ways (12)	years	ways
5	advantages (8)	women (11)	respects	cases
6	countries (5)	years (11)	countries	countries
7	parents (5)	Americans (8)	instances	teachers
8	reasons (5)	things (8)	areas	times
9	students (5)	reasons (8)	aspects	studies

10	programs (4)	students (7)	examples	women
11	disadvantages (4)	children (6)	teachers	children
12	companies (4)	problems (5)	times	programs
13	aspects (2)	changes (4)	schools	areas
14	games (2)	colleges (4)	parts	schools
15	times (2)	advantages (3)	species	problems
16		Britons (3)	women	states
17		criminals (3)	commentators	members
18		lives (3)	kinds	factors
19		men (3)	studies	things
20		parents (3)	factors	issues

The noun collocates of “many” in THAI can be classified into two major groups. The first one includes collocates whose occurrences on the list are largely attributed to the topics of the essays, i.e. “animals”, “companies”, “countries”, “games”, “parents”, “programs” and “students”. The other group consists of general or abstract nouns that are not specifically related to the contents of the essays in THAI, namely “advantages”, “people”, “things”, “ways”, “disadvantages”, “aspects”, reasons” and “times”. Based on this categorization criteria, it can be said that collocational patterns of “many” in THAI are similar to those in NATIVE as the latter also contains these two major groups of collocations, though it exhibits a wider range of noun collocates, especially in the group of general/ abstract nouns, e.g., “cases”, “years”, “changes” and “problems”, which do not appear on the THAI list. The BNC and COCA also display predominance of general/ abstract nouns as statistically significant collocates of “many”. This reflects a tendency in which this quantifier is used with general/ abstract nouns as observed by Dichtel (2016) and, interestingly, this is also found on the THAI list.

It is this group of nouns that suggest different usage patterns between “some” and “many” in THAI, as they are the type of nouns that are missing from the collocation list of “some” (Only one general noun, i.e. “people”, collocates with “some” in THAI). This phraseological discrepancy between “many” and “some” is linked to the two quantifiers’ different functional contributions to the students’ essays, as will be shown in 5.2.2 below.

5.2.2 Local textual functions of “many”

Based on the analysis of textual environment of the collocational patterns of “many”, it is found that the quantifier occurs significantly in THAI because they it helps carry out a variety of functions in the organization of students’ essays. They can be put into five groups, two of which are similar to those of “some” and are discussed first.

5.2.2.1 Introducing counter-arguments

Like “some”, “many” is also used to introduce counter-arguments or opposing views and this is also often realized through the similar core phraseological pattern: MANY + PEOPLE NOUN + SPEECH/ COGNITION VERB. However, “many” is slightly different in that it denotes a larger quantity than “some”, which indicates that the opposite views are held by a large number of people. This in turn suggests that the writer’s argument challenges a dominant view. In the example below, the student writer’s use of the word “still” shows his/ her emphasis on the argument, despite the fact that it goes against a popular view suggested by the reference to “many”.

I understand that animal experimentation might help human in curing diseases and many scientists are concerned about the problem but I still believe that using animals for test does not have efficiency enough. I think it is not worthy compared with animal’s life.

It must be noted that although “many” is also used to introduce counter-arguments, it does not tend to serve an additional purpose of structuring the essay in the way “some” does in THAI. This is reflected through the fact that out of 29 cases of the “MANY + PEOPLE NOUN + SPEECH/ COGNITION VERB” construction, only 7 of them (24.14%) occurs in a topic sentence that begins a new body paragraph. The other 22 occurrences were found to occur in the middle of a paragraph, suggesting that the reference to counterarguments serves to develop the main idea in each body paragraph, rather than creating an overall structure of an essay. Moreover, of these 22 cases, 8 instances (36.36%) occur in the introductory

paragraph, while the use of “some” for introducing counter-arguments appears only in a body paragraph. This may be because a reference to a large number of people who have differing views serves to highlight the importance of the issue under discussion given that the main purpose of an introductory paragraph in an argumentative essay is to provide contextualization of the topic and grab readers’ attention (Hyland 1990).

5.2.2.2 Supporting claims

Also like “some”, the collocation between “many” in THAI and the nouns shown in Table 2 are used to support or elaborate student writers’ claims or arguments. This function is mostly performed through “many” and its content-driven collocates indicated in 5.2.1 and often realized textually through its occurrence in a clause following the topic sentence of a body paragraph. Again, through its semantic property, the use of “many” can strengthen the supporting information, showing that entities relevant to a given topic are in a great number and variety. The paragraph below illustrates the use of “many” for this purpose.

Another reason why I believe the animal experimentation is wrong is that nowadays there are many innovations and they can be used for testing chemical or other substances. I think people should use these technologies to study and test the toxin, chemical structure, molecule of substance, etc. It is certainly safer than using animal experimentation and many animals do not have to receive the effects from testing or die anymore.

In the above paragraph, the reference to “many animals” can be seen as the student’s strategy to strengthen his/her claim about benefits of an animal-free approach to scientific experiment. It must be noted that in the sample paragraph this function also manifests explicitly through the co-occurrence of “many” with the intensifying words “certainly” and “anymore”.

These two functions of “many” are similar to those of “some” probably because both express meanings related to quantity of entities related to the essay topics and to the argumentative kind of writing. Also, the collocational patterns that realize these functions are shared by both quantifiers, i.e. “people” and content-driven

collocates. However, as discussed in 5.2.1, there are clusters that are specific to “many” only and this group of collocates have been found to contribute to local textual functions of “many”, which are different from those of “some” in THAI, as discussed below.

5.2.2.3 Signaling elaboration / list

The keyness of “many” in THAI is also attributed to its use for signaling elaboration. This function is often realized through the collocation between “many” and three general/ concept nouns, i.e. “things”, “reasons” and “advantages”. In many cases, these clusters are embedded in two textual positions: (1) in the thesis statement, i.e. the final sentence in the introductory paragraph leading to body paragraphs and (2) in the topic sentence of a body paragraph. A reference to “many” in these positions suggests to readers that quite a number of points relevant to the main idea are to be discussed in the body paragraphs that follow or within the body paragraph. The following sample extract is the first paragraph of an essay from THAI, with the cluster “many aspects” signposting arguments to be presented in the following body paragraphs:

Recently, there is a controversy about whether people should support factory farming or local farming. A side that supports local farming pinpoints so many disadvantages of factory farming such as a bad animal welfare which is a very sensitive topic throughout the world. On the other hand, factory farming supporters claim that local farming cannot replace factory farming to reduce the world hunger. However, although some people might argue that factory farming is the best solution to the world hunger, it should be banned because local farmin;, in fact, it is a more sustainable way to tackle the problem and it is significantly better than factory farming in many aspects.

Apart from their positions in the essays, this function is also manifested through the fact that clauses that contain the collocation of “many” and those nouns, giving a vague reference, are often followed by an illustration or a list of specific information, which helps elaborate what “many” refers to, as can be seen in the paragraph below.

*It is said that computer has incredibly a big impact on our everyday life. We use computer for **many reasons**: to work, to play games, to communicate with other people and many more. Basically, these are just the ways people use computer without any awareness.*

It must be noted that while the signaling-elaboration function of “many” in THAI is limited to its collocation with only a few general nouns, it is realized through much more varied general noun collocates of “many” in NATIVE, as illustrated below.

(23) *Men and women can compete against one another in **many things**. They can compete equally for a sells job or they can compete equally for a political position, but physically most men and women can not equally compete against one another.*

(24) *In his 10 years in office de Gaulle implemented **many changes** in the french political system. The traditional powers bestowed on him were ruthlessly exploited, used and abused. These included Head of State, Commander-in-Chief of Armed Forces and the power to appoint or dismiss the Prime Minister of his choice. De Gaulle was also given carte blanche to define European, domestic, defence and foreign policy and he did so, creating the office of the presidency and turning the presidency into the centre of decision-making. He became the chief policy maker and in the fields of foreign and defence policy he left his hallmark firmly implanted.*

5.2.2.4 Synthesizing points

Apart from signaling further development of ideas, the quantifier “many” in THAI is also used to refer to all the points that have been discussed in an essay. These points are captured, packaged together and referred to through the collocation between “many” and general/ concept nouns. In THAI this is often a case with the nouns “ways” and “advantages”. Not surprisingly, this usage pattern of “many” tends to be found in the concluding paragraph of an essay or in the last sentence of a

body paragraph. This usage is exemplified below, with (1) as a concluding paragraph and (2) the first body paragraph of an essay.

(25) *“In conclusion, there are many ways to evaluate how good the students are in certain subjects. The final exam is a big word for students and parents. If there is no final exam, parents would not expect too much from their children and students would not expect too much about their grade.”*

(26) *“First of all, computer makes your life easier. It has a lot of functions that can help our lives such as calculator, word processor, organization, etc. Previously, we had to recognize everything by ourselves or we may write it in our notebooks which may be lost anytime. When we have computers, everything seems easier. We can present a project in an easy, fun, interesting way. We cannot reject that computer help our lives in many ways.”*

It should be noted that while “many ways” is shared by THAI and NATIVE, it seems to be used for different textual functions. Native speaker learners tend to use it for signalling elaboration, rather than to synthesize all the points that have been discussed. For example:

The cellular telephone has changed people’s lives in many ways: the main way being the availability to be reached at any time. Communication is readily available in cars and on the streets now. People are able to use their car phones in case of an emergency; if they are stranded or broken down. They can take their cellular phones on vacations and to their friends houses if they are expecting an important call. To me, it is a bit ridiculous, but lots of people feel like they need this easy access to a phone.

An attention to local textual functions of “many” therefore helps to unearth subtle differences between native and Thai learners of English in terms of the textlinguistic dimension of “many”, which goes beyond its phraseological patterns.

5.2.2.5 Establishing significance of an essay topic

While the above four textual functions are more or less common in both THAI and NATIVE, there is one textual function of “many” that is particularly distinctive of argumentative essays written by Thai undergraduates. That is, “many” contributes to the ways in which Thai students establish the significance of the topic under discussion, whether by contextualizing, problematizing or even dramatizing a particular state of affairs relevant to an essay topic. This local textual function is formally achieved through a combination of particular lexicogrammatical and textual patterns of “many”. First, both general noun and content-driven collocates of “many” were used for this function. Secondly, the collocation occurs in the opening sentence of an essay. Finally, it tends to co-occur in close proximity with temporal expressions denoting the present time, e.g., “nowadays” or “now”. All these together express an idea that the noun referent is part of a current phenomenon that concerns a large number of people or entities. Beginning an essay with a statement that indicates relevance to the present time and a large quantity of something can thereby raise the profile of an issue under discussion. The sample paragraph below illustrates this function of “many” in THAI:

These days, many schools in Thailand are promoting themselves with classrooms full of computers and high technology devices. Technology becomes part of the ‘selling point’ that these private schools used to call out for more parents, or more customers. Many parents in this generation also take the access of computers in the classroom as part of their consideration to choose the best school for their children.

5.3 Every

The determiner “every” is often described in comparison with “all” as they are considered semantically equivalent except that “every” is distributive, i.e. it picks out the members of a set, while “all” considering them in mass (Quirk et al. 1985: 382). Therefore, the use of “every” is often described in terms of its colligation pattern that it is used to modify a singular noun. The analysis below provides a description of “every” in terms of its collocations and discourse functions.

5.3.1 Collocates of “every”

The number of nouns that are immediately modified by “every” and appear repeatedly in THAI is relatively low, compared with “some” and “many”, with a frequency of 67. As shown in Table 4 below, only nine noun types recur with “every” in THAI. Upon an examination of the whole cluster list, it was found that “every” collocates with a variety of nouns but most of them occur only once, e.g., “every product”, “every visit” and “every zoo”. Even some nouns which are related to contents of the essays and found in the collocation lists of the previous two quantifiers do not turn up in the list of “every”, including “student”, “program” and “school”. Table 4 below demonstrates the nouns that collocate with “every” more than once in the learner corpora.

Table 4. Top noun collocates of “every” in THAI, NATIVE, BNC and COCA

No.	THAI	NATIVE	BNC	COCA
1	time (4)	day (16)	day	day
2	day (3)	year (13)	case	year
3	life (3)	person (5)	months	time
4	animal (2)	time (5)	aspect	student
5	child (2)	American (4)	year	aspect
6	creature (2)	facet (4)	time	child
7	site (2)	human (4)	years	person
8	society (2)	individual (4)	stage	state
9		woman (4)	minutes	case
10		child (3)	person	effort
11		member (3)	man	level
12		case (3)	child	man
13		country (2)	member	week
14		man (2)	word	country
15		morning (2)	citizen	school
16		night (2)	bit	morning
17		part (2)	effort	night
18		teacher (2)	partner	word
19			citizen	step
20			level	month

It seems that collocational patterns of “every” in THAI and in the native speaker corpora are in a similar fashion in that “every” is used repeatedly with general nouns

that denote time, place, and person, some of which are related to the topics of the essays in the learner corpora, including “every child” (THAI) and “every American” (NATIVE). What is interesting is that temporal nouns are the most frequent collocates of “every” in all the corpora, with “time” and “day” occurring on all the four lists and the latter occupying the first place on the lists of the three native speaker corpora. Nevertheless, the native speaker corpora display a wider range of words denoting duration, e.g., “year”, “month”, “week” and “night”. A major point of difference between THAI and the three native speaker corpora is that such general nouns that denote abstract concepts as “facet”, “case” and “aspect” occur in the native speaker corpora but do not appear at all in THAI. Note that this is also a case with the collocation pattern of “some” in THAI as discussed in Section 5.1.

The other difference in usage pattern of “every” in THAI and NATIVE reflected through a concordance analysis deals with syntactic distribution patterns of “every” as shown in Table 5 below.

Table 5. Syntactic distribution pattern of “every” two-grams

Syntactic function	THAI (67)	NATIVE (141)
Subject	24 (35.82%)	42 (29.79%)
Object	9 (13.43%)	13 (9.22%)
Adverbial	33 (49.25%)	83 (58.87%)
Erroneous case	1 (1.49%)	3 (2.13%)

Based on the above table, it seems that Thai undergraduates tend to use “every” to quantify the subject and object in a clause more often than their native speaker counterparts while the latter go for the adverbial. Given that the subject of a clause refers to the agent or doer of an action and the object the affected, it can be said that “every” tends to be used by Thai students to highlight the number of entities involved in an action or event while native speaker learners often emphasize additional information about frequency in which an action occurs. This in a way corresponds to the fact that there is a wider array of time noun collocates in NATIVE as noted above. For instance,

THAI

Some people who support testing on animals said that when scientists do the experiments, they will try not to hurt or torture animals. I understand their point. However, every life is important. No one wants to suffer or die. Some scientists might avoid using painkillers. According to Knight, a member of the European veterinary specialist College of Animal 'Welfare Behavioral Medicine, "Although painkillers are sometimes given often they are not, partly owing to concerns this may affect the experiment's outcomes". (2012, para ll). This shows that many times the scientists avoid using painkillers to alleviate the pains of animals. They would feel every pain and suffering for a while, or they might die.

NATIVE

The above assertion from Daniel Callahan describes to the letter the abortion debate in American society. The disagreement has been clear for decades, and the two sides are speaking out louder and stronger every year. Each side has its own set of values and beliefs and the fact that those values conflict causes each side to rethink what they believe and hold onto those beliefs as tight as possible.

The different weight of information highlight through “every” is manifested more clearly when we look into local textual functions of the quantifier in the two learner corpora.

5.3.2 Local textual functions of “every”

It is found that “every” in THAI is often embedded in clauses that come after the topic sentence of a body paragraph, supporting the writer’s arguments. However, its claim-supporting function is phraseologically realized in a way that is different from the patterns found with “some” and “many”. The concordance analysis, illustrated in Figure 2 below, reveals that “every” tends to occur in textual environment where two types of lexical items are found:

- (27) expressions with an intensifying force, e.g., degree intensifiers, truth emphasizez and other quantifiers, e.g., “everything”, “everywhere”,

“most”, “no matter”, “even”, “own”; in some cases, the word “every” is repeated.

- (28) modal verbs that express a relatively strong degree of probability, commitment or obligation, including “must” and “have to”.

<p>of privacy protection is needed more among an beings want to change. If everything and arch <u>everything</u>, <u>everyplace</u>, <u>every story</u> and n on the Internet at <u>almost everywhere</u> and e to experiment instead of animals. ble and intelligent. Computers are set up at about the other like animals and trees. But pay for the expenses in Internet connection 's computer. Then the employer can control hen you can search <u>everything</u>, <u>everyplace</u>, ns of animals are used for experimentation what the students of my major <u>have to</u> face viate the pains of animals. They would feel 's life. Animal experimentation is risky for believe that single-sex education is <u>best</u> for</p>	<p>every generation in order to control our privacy every creature dies, human beings will live for every information <u>all around the world</u>. every time they like. For example, they can look Every species have their <u>own</u> hearts. We are every office <u>right now</u>. In our free time, we every single life <u>in the world</u> has the right to every month - causing <u>so much</u> money to their every employee's computer. In conclusion, if we every story and <u>every information all around the</u> every year for various studies and researches, every semester find after putting up with this every pain and suffering for a while, or they every animal that was used for experimentation. every child. The singe-sex education might</p>
---	--

Figure 2. Concordance lines of “every” in THAI

The word “every” itself articulates an intensifying force. Its collocation with time and place nouns, for example, indicates an event or action is frequent or common. The fact that it co-occurs with another expression with intensifying force in the above groups not only creates a cohesive chain of intensifying or emphatic expressions but also a hyperbolic effect in statements.

Such a strong intensification of the statements gives special reinforcement to the writer's arguments. In the examples from THAI below, the sentences in which “every” is embedded, follow the writers' explicit objection to the counterargument in (a), arguments about significance of other creatures and computers in (b) and (c), respectively.

- (29) *However, I find this argument rather weak and view this issue in a totally different way. In every society, no matter how small or big it is, “rules” are needed to keep it disciplined and prevent undesirable actions.*

- (30) *Moreover, laws can change if human beings want to change. If everything and every creature dies, human beings might live for a year or maybe more, but the world would not be beautiful anymore.*
- (31) *Computers respond to everyone's lifestyle. If you are a student, you can use it for your assignments because a computer has many programs for your assignment, for example, Microsoft word, power point or Microsoft excel. Or if you are a teacher or businessman, you can also use those programs. Moreover, you can connect the internet with your computer and then you can search everything, every place, every story and every information all around the world."*

This co-occurrence pattern of two-word clusters of "every" and intensifying expressions is also found in NATIVE, as shown below:

- (32) *About two years ago, another type of lenses were made accessible to the public. They're called "disposables" and I can't even begin to explain what a difference they make. You can wear them for a week and then throw them away -- with no worries about cleaning them and taking them out every night before going to sleep. It's such a great feeling to wake up in the morning, open your eyes, and be able to see perfectly! In the past, I'd have to "feel my way" to the bathroom until I could stumble upon my contact lense case and insert them with my nose up against the mirror.*
- (33) *Gathered around the tube, the American family learned the 'idea' way of life. But the picture of perfection was the same in every home. We all received the same information. We all saw the same families, The Cleavers and The Andersons, etc. telling how we should behave.*

Based on the above analysis, both Thai and native speaker learners seem to use "every" in the same way in that it serves to strengthen the writers' claims or arguments. However, an examination of distribution of "every" in the two corpora reveals a slightly different pattern as shown in Table 6 below.

Table 6. Distribution pattern of “every” in THAI and NATIVE

Corpus	Introduction	Body	Conclusion
THAI	7 (11.67%)	48 (80%)	5 (8.33%)
NATIVE	27 (19.15%)	101 (71.63%)	13 (9.22%)

As seen above, in both corpora, “every” is found the most in the body part of essays, where the writers present their arguments in support of their thesis statements. This is not surprising given that the body part is the longest part of an essay and that its major function is discussing grounds for thesis stated in the introduction (Hyland 1990). However, a pattern worth of note is found when comparing the distribution of “every” between THAI and NATIVE section by section. Thai undergraduates tend to use “every” less frequently than native speaker learners in the introduction but more often in the body part of their essays. Given that the body and introduction parts of an argumentative essay have different functions as explained in Section 4.3, a strong intensification created through “every” can be associated with the learners’ different preferred discourse strategies. Native speaker learners tend to use hyperbole not only to support their arguments but also to raise the significance of the topics while Thai students’ use of “every” centers around the generalization of their claims to back up the thesis. This observation, however, is based on small figures. Further studies can be conducted to explore this distribution pattern of functions of “every”

6. Discussion

The above applications of corpus linguistic concepts in the analysis of keyword quantifiers in learner corpora have demonstrated quite a few important points with regards to Thai students’ writing style when compared with native speaker learners’. First, in terms of phraeological patterns, the quantifiers are used with both similar and different types of noun collocates on the lists. Those shared by both corpora can be taken to provide evidence of usage profile of the quantifiers. Although quantifiers are often described as being associated with a particular grammatical category, they

in fact also co-select particular lexical items, as reflected by the fact that a certain words collocate with the quantifiers across the learner and general corpora, e.g., “every time”, “some people” and “many ways”. At the same time, quite a number of collocates of the three quantifiers in the two learner corpora are different. The disparity, on the one hand, reflects differences in essay topics in the corpora. However, some collocates point to the native-vs.-non-native stylistic distinction. The most remarkable case is perhaps general noun collocates denoting abstract concepts, e.g., “case (s)” and “kind (s)”. As revealed in 5.1, this group of nouns does not occur at all on the collocational lists of “some” and “every” in THAI whereas the corresponding lists of NATIVE, BNC and COCA have several of them in top places of the ranking. Although these nouns are found to collocate with “many” in THAI, the number of their types is relatively limited compared with NATIVE. This group of nouns is one of the central elements in academic writing, as pointed out by Jiang and Hyland (2016, 2017) and Flowerdew and Forest (2003), which refer to this noun group as “metadiscoursal noun” and “signaling noun”, respectively. Although the argumentative essay is arguably not exactly the same as typical academic writing, such as research abstracts or dissertation, which are studied in the above cited studies, it can be regarded as one of the first steps that prepare undergraduate students for academic writing. The absence of concept noun collocates with the quantifiers in THAI suggests that the writing by Thai students represented by this corpus suffers a lack of an important group of lexical items that can enable them to write academically in the future. This calls for a need to equip Thai and perhaps other non-native students with knowledge and skills in using the phraseology of “quantifier + abstract concept nouns” in their writing classes.

With respect to local textual functions, the analysis has shown that these quantifiers turn up as keywords in THAI because they perform one dominant function in the essays, shared by all the quantifiers under study, i.e. that of supporting claims. This reflects a strong tendency among Thai undergraduates to support their theses and arguments by drawing upon the concept of quantity. In other words, a reference, though vague, to a noticeable quantity of something is a dominant stylistic strategy in Thai undergraduates’ argumentation. While it can be argued that native speaker learners also use these quantifiers to support their claims, as revealed by the comparisons of their phraseological and textual patterns above, these quantifiers are not used significantly in NATIVE. This in turn suggests that

Thai learners' argumentation relies heavily on the concept of quantity, to the extent that a number of quantifiers become statistically significant in their writing, and that native speaker learners may employ other strategies in their essays. The repetition of a group of quantifiers for the same purpose thereby creates repetitiveness in the rhetorics of Thai students' English argumentative writing.

The problem of repetition in Thai EFL learners' essays is manifested not only through repeated discourse strategy but also through the fact that a particular local textual function is realized through repeated phraseological patterns. This is demonstrated through the above analyses of the relationship between collocations and textual functions of those quantifiers. The case of "some" and "many" and their shared local textual function of "introducing counterarguments", for example, can illustrate this point. While the function is realized lexicogrammatically in more various forms in NATIVE, the phraseological expressions that articulate the same function in THAI are limited to the core pattern [SOME / MANY + PEOPLE NOUN + SPEECH/ COGNITION VERB]. This may be attributed to the influence of teaching that emphasizes certain expressions and, at the same time, reflect Thai students' learning strategy of memorizing taught expressions.

Apart from shared local textual functions, the three quantifiers have been found to have their own distinctive functions in the essays. For example, the keyword "many" is mainly used not only for introducing counter-arguments and supporting claims like "some" but also for summarizing and signaling elaboration. Moreover, even those with similar textual properties seem to demonstrate variation in the ways they are used in texts. The counterargument-introducing function of "some" and "many", for instance, tends to occur at different positions in the essays, with the former often starting a body paragraph and contributing to the structure of an entire essay. Also, the quantifier "every", though also used to provide support to the writers' positions, fulfills the function through creation of hyperbolic claims while the other two often do through introducing examples or causes/ effects. It must also be noted that such variations among the functional profiles of the quantifiers are in many cases realized through their co-textual features and textual positions. All these yield support to a recent observation in corpus linguistics that lexis, grammar and text are inextricably linked, i.e. that meaning in English is expressed not only through lexicogrammatical patterns but also through textual positions in which these patterns are embedded, and that different lexicogrammatical patterns of a lexical item

are linked to its idiosyncratic textual properties (Hoey 2005; Hoey and O'Donnell 2015).

The interesting point is that this linguistic mechanism of English is also exhibited in learner data. This corpus-informed evidence can perhaps be linked to theory in interlanguage studies that learners have their own idiosyncratic linguistic properties and, as the present study has suggested, these features not only relate to grammatical aspects, but also to discoursal features of such grammatical words as quantifiers.

7. Conclusion

This paper has adopted a corpus-driven approach to Thai undergraduates' writing by starting from three different keywords in a corpus of their essays and extending the analysis from the lexical to phraseological and textual levels. On the one hand, this study can be seen as another learner corpus research project that discusses similarities and differences between non-native and native speaker learners' English writing, focusing on quantifiers, which have rarely been analysed in learner corpus research. While quantifiers are often dealt with in the EFL context as grammatical items that require particular types of nouns and hardly ever in the writing class, the present study has shown that these small words play an important role in Thai EFL students' English writing when compared with native speaker learners. They are one of the major linguistic categories that accounts for rhetorical similarities and differences between Thai and non-native learners' essays. Their patterns of distributions and co-occurrences are shown to play a major part in textual structure, organization, rhetorics of argumentation as well as creation of meanings in the essays written by both groups of learners. Also, the significant recurrence of these items in Thai students' essays point to the problem of repetitiveness in their writing, whether at the lexical, phraseological or discourse levels. This clearly has pedagogical implications. More attention to stylistic and discourse strategies are needed in a writing class, in addition to the already-existing emphasis on the macro structure of essays, grammar and expressions. In fact, a theoretical perspective on the relationship between linguistic forms and their textual functions should be more or less integrated into the design and development of English instruction and materials to promote naturalness in EFL writing. In other words, such usage-based findings and theoretical

concepts can serve as input to influence the development of learners' interlanguage.

On the other hand, although this study is applied linguistic in nature, i.e. it applies theoretical concepts on the relationship between lexicogrammatical patterns and textlinguistics to a description of quantifiers in learners' writing, it might be a good idea to look at the findings and relate them back to the theoretical description of the quantifiers and to the perspective the study draws on. Although the text-lexicogrammar relationship has been increasingly addressed in corpus linguistics, they are mostly illustrated by studies of "standard" English texts, not by data produced by learners. The fact that there are similarities between the phraseological and text-functional profiles of those quantifiers in THAI and NATIVE lend support to those theories on the text-lexicogrammar relationship like Mahlberg's (2013) and Hoey's (2005) that a lexical item has a tendency to occur in particular lexicogrammatical patterns and textual environment. As shown through the analyses, even data from learners from different background display a tendency in which quantifiers occur in particular phraseological and textual environment. It is possible that this textual dimension of the given quantifiers are specific to the argumentative essay even though the texts are different in their topics and their writers have different proficiency levels. In other words, the identified textual functions of these quantifiers are perhaps "local" to learner argumentative essays. A further study might be conducted on other text types or on general corpora to see whether these quantifiers perform similar sorts of textual functions. Corpus linguistics thereby offers not only techniques but theoretical concepts that can benefit applied linguistics while findings from the application can bring us back to confirm, refute or extend the scope of the theoretical descriptions of the English language.

References

- Anthony, Laurence. 2014. AntConc (Version 3.4.3) [Computer Software]. Tokyo, Japan: Waseda University. Available from [<http://www.laurenceanthony.net/>].
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman grammar of spoken and written English*. London; New York: Longman.
- Bondi, Marina. 2008. Emphatics in academic discourse: Integrating corpus and discourse tools in the study of cross-disciplinary variation. In Annelie Ädel and Randi Rappen

- (eds.) *Corpora and Discourse: The challenges of different settings*, 31-56. Amsterdam: John Benjamins
- Chen, Yu-Hua and Paul Baker. 2010. Lexical bundles in L1 and L2 academic writing. *Language Learning and Technology* 14(2): 30-49.
- Chu, Chia-Ying, Alison Gabriele, and Utako Minai. 2014. Acquisition of quantifier scope interpretation by Chinese-speaking learners of English. In Chia-Ying Chu, Caitlin E. Coughlin, Beatriz Lopez Prego, Utako Minai, and Annie Tremblay (eds.), *Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012)*, 157-168. Somerville, MA: Cascadia Proceedings Project.
- Conrad, Susan and Douglas Biber. 2005. The frequency and use of lexical bundles in conversation and academic prose. In Wolfgang Teubert and Michaela Mahlberg (eds.), *The corpus approach to lexicography*, 56-71.
- Dichtel, Frédéric. 2016. A quantifier used on *many* occasions: *Many* evoking diversity in positive sentences. *International Journal of Corpus Linguistics* 21(1): 80-104.
- Flowerdew, John. 2002. *Academic Discourse*. London: Longman.
- Flowerdew, Lynne. 2008. *Corpus-based analyses of the problem-solution pattern: A phraseological approach*. Amsterdam: John Benjamins.
- Flowerdew, John and Richard W. Forest. 2003. *Signalling nouns in English: A corpus-based discourse approach*. Amsterdam: John Benjamins.
- Granger, Sylviane. 2015. Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research* 1(1): 7-24.
- Granger, Sylviane and Magali Paquot. 2010. Lexical verbs in academic discourse: A corpus-driven study of learner use. In Maggie Charles, Susan Hunston, and Diane Pecorari (eds.), *Academic Writing: At the Interface of Corpus and Discourse*. London: Bloomsbury.
- Halliday, Michael. 1985. *An introduction to functional grammar*. London: Routledge.
- Hinkel, Eli. 2002. *Second language writers' text: Linguistic and rhetorical features*. Mahwah, NJ : Lawrence Erlbaum Associates.
- Hoey, Michael. 2005. *Lexical priming: A new theory of words and language*. Stroud: Taylor; Francis.
- Hoey, Michael and Matthew Brook O'Donnell. 2015. Examining associations between lexis and textual position in hard news stories, or according to a study by. In Nicholas Groom, Maggie Charles and Suganthi John (eds.), *Corpora, Grammar and Discourse*. Amsterdam; Philadelphia: John Benjamins.
- Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press.
- Hyland, Ken. 1990. A genre description of the argumentative essay. *RELIC Journal* 21:

66-78.

- Hyland, Ken. 2003. *Second language writing*. Cambridge: Cambridge University Press.
- Ionin, Tania, Tatiana Luchkina, and Anastasia Stoops. 2014. Quantifier Scopepe and Scrambling in the Second Language Acquisition of Russian. In Chia-Ying Chu, Caitlin E. Coughlin, Beatriz Lopez Prego, Utako Minai, and Annie Tremblay (eds.), *Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012)*, 181-190. Somerville, MA: Cascadilla Proceedings Project.
- Katsos, Napoleon, Chris Cummins, Maria-José Ezeizabarrena, Anna Gavarró, Jelena Kuvač Kraljević, Gordana Hrzica, and Kleanthes K. Grohmann et al. 2016. Cross-linguistic patterns in the acquisition of quantifiers. *Proceedings of the National Academy of Sciences* 113(33): 9244-9249.
- Jiang, Feng Kevin and Ken Hyland. 2016. Nouns and academic interactions: A neglected feature of metadiscourse. *Applied Linguistics*, 1-25.
- Jiang, Feng Kevin and Ken Hyland. 2017. Metadiscursive nouns: Interaction and cohesion in abstract moves. *English for Specific Purposes* 46(1): 1-14.
- Leech, Geoffrey, Paul Rayson, and Andrew Wilson. 2001. *Word frequencies in written and spoken English: Based on the British National Corpus*. Longman: London.
- Li, Jie and Norbert Schmitt. 2010. The development of collocation use in academic texts by advanced L2 learners: A multiple case study approach. In David Wood (ed.), *Perspectives on formulaic language: Acquisition and communication*, 2-46. London; New York: Continuum.
- Lima, Suzi. 2014. The acquisition of the count/mass distinction in Yudja (Tupi): Quantifying 'quantity' and 'number'. In Chia-Ying Chu, Caitlin E. Coughlin, Beatriz Lopez Prego, Utako Minai, and Annie Tremblay (eds.), *Selected Proceedings of the 5th Conference on Generative Approaches to Language Acquisition North America (GALANA 2012)*, 181-190. Somerville, MA: Cascadilla Proceedings Project.
- Lorenz, Gunter R. 1999. *Adjective intensification - Learners versus native speakers. A corpus study of argumentative writing*. Amsterdam; Atlanta, GA: Rodopi.
- Mahlberg, Michaela. 2003. The textlinguistic dimension of corpus linguistics: The support function of English general nouns and its theoretical implications. *International Journal of Corpus Linguistics* 8(1): 97-108.
- Mahlberg, Michaela. 2005. *English general nouns: A corpus theoretical approach*. Amsterdam: John Benjamins.
- Mahlberg, Michaela. 2007. Clusters, key clusters and local textual functions in Dickens. *Corpora* 2(1): 1-31.
- Mahlberg, Michaela. 2009. Local textual functions of *move* in newspaper story patterns. In Ute Römer and Rainer Schulze (eds.) *Exploring the Lexis-Grammar Interface*, 265-287.

- Amsterdam: John Benjamins.
- Mahlberg, Michaela. 2013. *Corpus stylistics and Dickens's fiction*. London: Routledge.
- Marsden, Heather. 2009. Distributive quantifier scope in English-Japanese and Korean-Japanese interlanguage. *Language Acquisition* 16(3): 135-177.
- McEnery, Tony and Andrew Wilson. 2001. *Corpus linguistics: An introduction*. Edinburgh: Edinburgh UP.
- Mukherjee, Joybrato. 2005. The native speaker is alive and kicking - linguistic and language-pedagogical perspectives. *Anglistik* 16(2): 7-23.
- Nesselhauf, Nadja. 2005. *Collocations in a learner corpus*. Amsterdam; Philadelphia: John Benjamins.
- Paquot, Magali. 2010. *Academic vocabulary in learner writing: From extraction to analysis*. London: Continuum.
- Paquot, Magali and Sylviane Granger. 2012. Formulaic language in learner corpora. *Annual Review of Applied Linguistics* 32: 130-149.
- Partee, Barbara. 1989. Many quantifiers. In J. Powers and K. de Jong (eds.), *ESCOL 89: Proceedings of the Eastern States Conference on Linguistics*. Papers presented at The Eastern States Conference on Linguistics, University of Delaware, Newark, 6-8 October. Columbus: OSU, 383-402.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London; New York: Longman.
- Rayson, Paul, Damon Berridge, and Brian Francis. 2004. Extending the Cochran rule for te comparison of word frequencies between corpora. In G. Purnelle, C. Fairon, A. Dister (eds.) *Proceedings of the 7th International Conference on Statistical Analysis of Textual Data (JADT 2004) Volume II*. Louvain-la-Neuve, Belgium, March 10-12, 2004, Presses universitaires de Louvain, 926-936.
- Rayson, Paul Edward. 2003. *Matrix: A statistical method and software tool for linguistic analysis through corpus comparison*. PhD Dissertation, Lancaster University.
- Sardinha, Tony Berber. 2004. *Linguística de corpus*. São Paulo: Manole.
- Scott, Mike and Christopher Tribble. 2006. *Textual patterns: Key words and corpus analysis in language education*. Amsterdam: John Benjamins.
- Stubbs, Michael. 2007. Quantitative data on multi-word sequences in English: The case of the word world. In Michael Hoey, Michaela Mahlberg, Michael Stubbs, and Wolfgang Teubert (eds.), *Text, Discourse and Corpora: Theory and Analysis*. London: Bloomsbury.
- Stubbs, Michael. 2015. The textual functions of lexis. In Nicholas Groom, Maggie Charles, and Suganthi John (eds.), *Corpora, Grammar and Discourse*. Amsterdam; Philadelphia: Benjamins.
- Tognini-Bonelli, Elena. 2001. *Corpus linguistics at work*. Amsterdam: John Benjamins.

Collocations and local textual functions of quantifiers in learner English essays 45

Wei, Yaoyu and Lei Lei. 2011. Lexical bundles in the academic writing of advanced Chinese EFL learners. *RELIC Journal* 42(2): 155-166.

Appendix 1

Keywords in THAI, when compared with NATIVE

No.	Word	Frequency in THAI	Per cent	Frequency in NATIVE	Per cent	LL value
1	students	339	0.84	222	0.09	654.25
2	single-sex	139	0.35	0	0	549.97
3	computer	185	0.46	55	0.02	489.98
4	online	117	0.29	0	0	462.92
5	computers	142	0.35	57	0.02	340.43
6	animal	110	0.27	19	0.01	333.05
7	privacy	97	0.24	12	0	311.78
8	information	143	0.36	85	0.03	289.94
9	animals	115	0.29	41	0.02	287.51
10	experimentation	75	0.19	2	0	278.79
11	internet	72	0.18	1	0	274.61
12	can	440	1.1	988	0.39	271.2
13	boys	89	0.22	19	0.01	257.32
14	exam	75	0.19	6	0	255.76
15	study	105	0.26	49	0.02	237.38
16	use	192	0.48	260	0.1	220.74
17	might	111	0.28	70	0.03	218.47
18	facebook	54	0.13	0	0	213.66
19	girls	86	0.21	32	0.01	211.87
20	moreover	62	0.15	6	0	206.51
21	education	134	0.33	128	0.05	205.22
22	social	139	0.35	149	0.06	195.42
23	class	82	0.2	34	0.01	194.23
24	some	225	0.56	402	0.16	191.36
25	studying	57	0.14	10	0	172.03
26	nowadays	50	0.12	6	0	161.48
27	technology	70	0.17	36	0.01	151.83
28	you	222	0.55	473	0.19	148.44
29	people	404	1.01	1229	0.49	137.16
30	do	261	0.65	662	0.26	130.36
31	think	111	0.28	148	0.06	129.5
32	work	117	0.29	168	0.07	127.02
33	learning	56	0.14	27	0.01	124.89
34	they	496	1.23	1700	0.68	122.23
35	schools	82	0.2	86	0.03	117.24
36	using	70	0.17	67	0.03	107.05
37	your	81	0.2	95	0.04	105.89

38	school	92	0.23	125	0.05	105.43
39	will	288	0.72	851	0.34	104.76
40	we	273	0.68	795	0.32	102.68
41	have_to	100	0.25	162	0.06	95.49
42	their	349	0.87	1170	0.47	91.74
43	friends	50	0.12	40	0.02	86.09
44	media	60	0.15	66	0.03	82.66
45	them	171	0.43	479	0.19	70.05
46	performance	28	0.07	11	0	67.66
47	good	103	0.26	224	0.09	66.75
48	or	273	0.68	941	0.38	66.14
49	learn	51	0.13	62	0.02	64.67
50	i	233	0.58	768	0.31	64.24
51	human	77	0.19	143	0.06	62.35
52	according_to	67	0.17	113	0.05	61.09
53	make	121	0.3	306	0.12	60.77
54	time	110	0.27	268	0.11	59.11
55	it	552	1.37	2399	0.96	53.92
56	are	470	1.17	1992	0.8	52.05
57	many	178	0.44	577	0.23	51.37
58	more	232	0.58	831	0.33	49.85
59	about	143	0.36	432	0.17	49.37
60	every	67	0.16	125	0.05	48.47
61	for	507	1.26	2224	0.89	47.16
62	such_as	69	0.17	155	0.06	42.5
63	children	107	0.27	330	0.13	35.12
64	easy	34	0.08	52	0.02	34.58
65	get	77	0.19	207	0.08	34.36
66	better	65	0.16	160	0.06	34.29
67	so	123	0.31	417	0.17	31.3
68	not	427	1.06	1971	0.79	29.58
69	than	96	0.24	305	0.12	29.22
70	know	63	0.16	168	0.07	28.57
71	n't	111	0.28	382	0.15	27.02
72	world	92	0.23	303	0.12	25.42
73	have	303	0.75	1363	0.54	24.52
74	should	150	0.37	581	0.23	24.46
75	want	62	0.15	179	0.07	23.78
76	new	59	0.15	175	0.07	21.27
77	like	80	0.2	277	0.11	19.12
78	different	58	0.14	181	0.07	18.49
79	our	119	0.3	470	0.19	17.97

80	us	58	0.14	185	0.07	17.47
81	my	61	0.15	200	0.08	17.07
82	because	139	0.35	589	0.24	15.41

Appendix 2

Content keywords in THAI, when compared with NATIVE

No.	Word	Frequency in THAI	Per cent	Frequency in NATIVE	Per cent	LL value
1	students	339	0.84	222	0.09	654.25
2	single-sex	139	0.35	0	0	549.97
3	computer	185	0.46	55	0.02	489.98
4	online	117	0.29	0	0	462.92
5	computers	142	0.35	57	0.02	340.43
6	animal	110	0.27	19	0.01	333.05
7	privacy	97	0.24	12	0	311.78
8	information	143	0.36	85	0.03	289.94
9	animals	115	0.29	41	0.02	287.51
10	experimentation	75	0.19	2	0	278.79
11	internet	72	0.18	1	0	274.61
12	boys	89	0.22	19	0.01	257.32
13	exam	75	0.19	6	0	255.76
14	study	105	0.26	49	0.02	237.38
15	use	192	0.48	260	0.1	220.74
16	facebook	54	0.13	0	0	213.66
17	girls	86	0.21	32	0.01	211.87
18	moreover	62	0.15	6	0	206.51
19	education	134	0.33	128	0.05	205.22
20	social	139	0.35	149	0.06	195.42
21	class	82	0.2	34	0.01	194.23
22	studying	57	0.14	10	0	172.03
23	nowadays	50	0.12	6	0	161.48
24	technology	70	0.17	36	0.01	151.83
25	people	404	1.01	1229	0.49	137.16
26	think	111	0.28	148	0.06	129.5
27	work	117	0.29	168	0.07	127.02
28	learning	56	0.14	27	0.01	124.89
29	schools	82	0.2	86	0.03	117.24
30	using	70	0.17	67	0.03	107.05
31	school	92	0.23	125	0.05	105.43
32	friends	50	0.12	40	0.02	86.09

33	media	60	0.15	66	0.03	82.66
34	performance	28	0.07	11	0	67.66
35	good	103	0.26	224	0.09	66.75
36	learn	51	0.13	62	0.02	64.67
37	human	77	0.19	143	0.06	62.35
38	make	121	0.3	306	0.12	60.77
39	time	110	0.27	268	0.11	59.11
40	children	107	0.27	330	0.13	35.12
41	easy	34	0.08	52	0.02	34.58
42	get	77	0.19	207	0.08	34.36
43	better	65	0.16	160	0.06	34.29
44	not	427	1.06	1971	0.79	29.58
45	know	63	0.16	168	0.07	28.57
46	n't	111	0.28	382	0.15	27.02
47	world	92	0.23	303	0.12	25.42
48	want	62	0.15	179	0.07	23.78
49	new	59	0.15	175	0.07	21.27
50	like	80	0.2	277	0.11	19.12
51	different	58	0.14	181	0.07	18.49

Raksangob Wijitsopon

Chulalongkorn University

Department of English

254 Phayathai Rd, Wang Mai, Khet Pathum Wan, Krung Thep Maha Nakhon 10330, Thailand

E-mail: rwijitsopon@gmail.com

Received: 2017. 01. 31.

Revised: 2017. 03. 29.

Accepted: 2017. 03. 29.