

A Systemic Functional Study of Thematic Organization in the English Writing of Korean College Students

Yerahm Choung*

(Seoul National University)

Sun-Young Oh

(Seoul National University)

Choung, Yerahm, & Oh, Sun-Young. (2017). A systemic functional of thematic organization in the English writing of Korean college students. *English Teaching*, 72(3), 119-144.

Employing Halliday's (1994) Systemic Functional Grammar approach, the present study examines the Korean EFL college students' way of organizing thematic structures in written English texts, with special focus on the influence of learner proficiency. A total of 150 learners' argumentative essays consisting of 50 essays each from three levels of learner groups were compared with 50 essays from its native counterparts, with inquiry focused on the thematic selection, usage of various theme features, and semantic distribution of topical themes. The results show that textual themes, multiple themes, and first person pronouns as topical themes are highly overused in the learner texts compared to the native essays. Overall, the level of proficiency significantly affects learners' choices regarding thematic structures although the enhanced proficiency does not involve the improvement in all aspects of thematic organization. From these findings, it is suggested that the EFL writing pedagogy should pay greater attention to the effective use of thematic organization as a means of enhancing textual coherence and embrace level-specific instructions.

Key words: Theme, thematic structure, systemic functional grammar, second language (L2) writing, argumentative writing, corpus

1. INTRODUCTION

Inspired by M. A. K. Halliday's (1994) work on Functional Grammar, a number of linguists have devoted substantial research effort to the role of the initial position of a clause in the distribution of ideas or development of the texts (Fries, 1995; Ghadessy &

* Yerahm Choung: First author; Sun-Young Oh: Corresponding author

Gao, 2000; Gosden, 1992; Green et al., 2000; Hewings, 2004; Jalilifar, 2010; Martinez, 2003; North, 2005). From this viewpoint, a sentence can be divided into 'theme' and 'rheme'; 'theme' refers to the first element in the sentence that represents "the starting point for the message" (Halliday & Matthiessen, 2013, p. 38) while 'rheme' is the component that comes after the theme and constitutes the remainder of the message. The semantic organization that the theme and rheme construct together is called a 'thematic structure.' In the thematic structure, the theme plays a prominent role in a clause, making a connection between the previous message and the following sentence by continuing the information provided by the preceding sentence. In turn, the rheme develops and elaborates the information or idea initiated in the theme. By analyzing the thematic structure of a text, it is possible to understand how the writer realizes the nature of his/her underlying concern (Halliday & Matthiessen, 2013, p. 133). The thematic structure is in close relationship with the distribution of ideas in the text and a writer's selection of theme pertains to the rhetorical purposes that the writer has in mind. Diverse theme types and their contents can thus be regarded as an important linguistic option for logical and coherent writing. Skillful writers are able to organize the theme and rheme in each clause constructing a coherent message.

Given its pivotal role in constructing information structure and maintaining textual coherence, a number of systemic functional researchers explored the writers' choice of thematic structure. Numerous studies focused on distinct groups of writers (Ghadessy, 1999; North, 2005; Wang, 2007) or specific types of writing (Fontaine & Kodratoff, 2011; Francis, 1990; Gómez, 1994; Hewings, 2004) as well as commonalities and/or disparities between different types of writers or genres of writing (Gosden, 1992; Jalilifar, 2009, 2010; Kim, 2012; Lee, 2007, 2008; Lu, 2002; Martinez, 2003; Nam & Park, 2015; Park & Nam, 2015). It was mostly agreed that less competent writers are prone to employ ineffective thematic strategy, using too simple a theme or too many themes at a time, and/or overusing a specific thematic structure such as existential *there* (Bloor & Bloor, 1992; North, 2005).

Studies conducted in the second language (L2) context have reported that L2 writers are likely to have difficulty in choosing appropriate thematic structure (Fontaine & Kodratoff, 2011; Hawes & Thomas, 1997; Lee, 2008; Mauranen, 1996). Comparative studies between L2 learners and English natives found that the learners' improper use of theme tends to result in weakened arguments. In order to understand the challenges that the non-native writers face with regard to theme selection and the specific changes that are required for them to overcome the issues, learner performance needs to be analyzed in careful consideration of their level of proficiency as those challenges and changes are likely to vary depending on the proficiency level. To compare the learners' use of theme across proficiency levels will allow us to see whether and how L2 writers' ability to construct coherent progression of ideas improves along with their competence development.

In addition, there is a clear need for statistical generalization as the preceding literature

has been based mostly on dozens of example essays (e.g., Hewings, 2004; Kim, 2012; Lee, 2008). Despite the considerable efforts made in this field thus far, research employing large-scale text corpora with statistical verification is still in its infancy. Recently, a few studies (e.g., Nam & Park, 2015; Park & Lu, 2015; Park & Nam, 2015) started to adopt computer-aided theme analysis which enabled a large-scale investigation of thematic choice. Most of these analyses, however, have not taken learner proficiency into account as a variable and limited to some subsets of theme features (e.g., theme length, theme function, and theme role). As an extension of the previous efforts, the present study aims to conduct a more comprehensive analysis of thematic structures with consideration of learner proficiency.

Employing Halliday's systemic functional approach, the present research focuses on the Korean EFL college students' construction of the thematic structure in English writing, with special focus on the effect of learner proficiency. The study uses corpora consisting of argumentative essays written by Korean EFL college students and native English speakers, and investigates the thematic selection and usages of distinctive theme features that have been broadly adopted in a number of systemic functional research studies (Gosden, 1992; Hewings, 2004; Jalilifar, 2009; Lee, 2007, 2008; Martinez, 2003; North, 2005; Park & Lu 2015; Park & Nam 2015). The theme features analyzed in the study include theme length and theme function such as textual, interpersonal, topical theme, and markedness of theme. In addition, thematising structures, which have been relatively neglected in previous studies, are also analyzed. It is worth examining them since the use of specific thematising structures may guide the reader towards a particular pattern of emphasis (Halliday & Matthiessen, 2013). Topical theme, which is an indispensable element in constructing a thematic structure, is also one of the main concerns of the present study. For a detailed analysis on the experiential content of theme, the current research investigates the distribution of semantic categories of the topical theme (Francis, 1990). By carefully analyzing the thematic structure of many thousands of instances, the current research aims to supplement previous literature and gain a better understanding of how learners employ theme and rheme in actual writing contexts.

The present study specifically seeks to answer the following research questions:

- 1) What are the differences in the use of various theme features among Korean EFL writers of different proficiency levels and native English writers?
- 2) What are the differences in the use of thematising structures among Korean EFL writers of different proficiency levels and native English writers?
- 3) What are the differences in the semantic distribution of topical theme among Korean EFL writers of different proficiency levels and native English writers?

2. REVIEW OF THE LITERATURE

2.1. Theme in Systemic Functional Grammar

Since Mathesius articulated the concept of a theme in the early 1930s, scholars of the Prague School (Fries, 1995) have constantly elaborated the notion of the theme and the thematic structure. Influenced by such work, Halliday adopted this concept into his Systemic Functional Grammar (SFG) framework. Halliday's SFG focuses on theme and rheme as fundamental elements for constructing a clause. In this framework, an element that functions as a starting point of a message is called a theme. The theme continues the information provided in previous sentences by "orienting the clause within its context" (Halliday & Matthiessen, 2013, p. 66). The rheme, on the other hand, comes after the theme by inheriting the message conveyed by the theme. The following examples, which are taken from Halliday and Matthiessen (2013, p. 90), display the way a clause is segmented into the theme and the rheme.

Theme	Rheme
<i>The duke</i>	<i>has given my aunt that teapot</i>
<i>My aunt</i>	<i>has been given that teapot by the duke</i>
<i>The teapot</i>	<i>the duke has given to my aunt</i>

The three sentences above consist of the same components; the elements, however, could be developed differently as described in the examples. What makes the three sentences different is having different starting points, in other words, selecting a different theme for the message. The difference between the first and the second sentence is whether the doer of the act (*The duke*) or the receiver of the action (*My aunt*) is functioning as a theme. When *by the duke* is introduced in the rheme of the sentence as in the second instance, *the duke* then becomes available as a natural starting point for the following sentence. In the third sentence, the theme *The teapot* could be read as indicating that the teapot may be the main topic of the clause. In this way, a certain choice of theme reflects a writer's intention to maintain coherence in the text.

The comparison among the variations suggests that although each clause refers to the same state of affairs, the different choice of theme contributes to a different meaning (Thompson, 2004). English proposes a great variety of semantically comparable propositions with dissimilar sentence structures; a writer can therefore change the starting point by moving the components of a clause around, considering which arrangement is more appropriate in a given context.

A theme can be given either 'unmarked' or 'marked' status, depending on whether the

choice of a theme is typical or unusual to occur in the thematic position of that mood class (Eggins, 2004). A theme position may be occupied by a mood structure constituent that typically occurs in the first position in the clauses of a given mood system (i.e., an unmarked theme), or by any other constituent from that mood class (i.e., a marked theme). For example, in a declarative sentence where the subject as a theme is regarded as a natural choice, thematic choices such as an adjunct or a complement are classified as ‘marked themes.’ The same principle can be applied to imperatives and interrogatives. A finite in an interrogative, a predicator in an imperative, and a *wh*-element in a *wh*-interrogative play marked roles. If there is any proper reason for choosing a theme other than the unmarked theme, a marked theme is used in the text (Halliday, 1994). Marked choices inevitably function as the writer-selected signals, as Gosden (1992) put it. It has been found that marked themes are employed more frequently in advanced writing (Hawes & Thomas, 1997; Jalilifar, 2010; Lee, 2008; Martinez, 2003; Park & Lu, 2015; Ping, 2007).

Theme markedness has been illustrated above in a single clause of a constituent. In addition, there are ‘multi-element constituents’ as a theme, which have been named as ‘thematising structures’ (Eggins, 2004; Halliday, 1994; Thompson, 2004). Thematising structures including ‘thematic equatives,’ ‘predicated theme,’ ‘thematised comment,’ and ‘preposed theme’ also contain unmarked and marked forms (see Table 3 for examples). The thematising structures show the speaker’s thematic selections related to the structures’ communicative functions.

Halliday (1994) labels the ideational thematic content ‘topical theme,’ arguing that the topical theme closely corresponds to what is generally called a ‘topic’. The topical theme is a vital element in the thematic structure, as every theme contains one and only one topical theme. Given its special status, topical theme has been analyzed in terms of semantic properties. Francis (1990) suggested ‘people/group/institutions,’ ‘concrete things,’ ‘abstractions,’ and ‘time expressions’ as the basic semantic categories of topical theme, and her category has been employed in several theme analyses with some modifications (Fries, 1995; Ghadessay & Gao, 2000; Hawes & Thomas, 1997; Lee, 2008; Vande Kopple, 1991). Lee (2008), for example, elaborated Francis’s (1990) classification into the following categories: (1) the first/second person, (2) animate and inanimate text participants, (3) global theme, meaning ‘the most central referent, often manifested in the writing topic,’ (4) abstractions vs. concrete things, (5) anaphoric nominalization, (6) text reference, which refers to a certain portion of text, and (7) locative circumstances.

Any element that precedes the topical theme can also be part of the theme. The preceding constituents have either textual or interpersonal functions. ‘Textual themes’ include conjunctions, conjunctive adjuncts, continuatives, and discourse markers (*well, now, yes, no, etc.*) as these elements signal the beginning of a message (Thompson, 2004). An ‘interpersonal theme’ expresses a speaker’s mood or attitude toward the message through

modal adjuncts, vocatives, and finite verbal operators. A theme belongs to one of the two categories, 'simple' or 'multiple,' depending on whether a theme is comprised of a topical theme only (i.e., a 'simple theme'), or additionally contains either a textual or an interpersonal constituent or both (i.e., 'multiple theme').

2.2. Systemic Functional Grammar Research

2.2.1. Studies on thematic structure

Early studies in SFG explored how writers employ thematic structure in their texts. These studies focused on specific groups of writers or particular types of writing (Fries, 1983, 1995; Ghadessy, 1999; Ghadessy & Gao, 2000; Gómez, 1994; North, 2005; Vande Kopple, 1991; Wang, 2007) or the comparison between various types of writers or genres of writing (Gosden, 1992; Jalilifar, 2009, 2010; Kim, 2012; Lee, 2007, 2008; Lu, 2002; Martinez, 2003; Nam & Park, 2015; North, 2005; Park & Nam, 2015). Multiple studies have found that novice/non-native and expert/native writers differ vastly in their choices of thematic structure.

The first noticeable difference has been reported in the use of theme function. A textual theme is one of the prominent features that shows differences according to the level of competence (Bolton et al., 2002; Green et al., 2000; Lee, 2007). It has been generally found that less proficient writers have a tendency to overuse textual themes, which results in disturbing continuity (Granger & Tyson, 1996; Green et al., 2000; Jalilifar, 2009; Park & Nam, 2015). Lee (2007) shows that unskilled writers significantly depend on the textual theme, in contrast to native writers who employ a minimal amount of the textual device in academic writing. This finding may be partially explained by Hawes and Thomas's (1997) suggestion that a textual theme functions as the "obvious player" in connecting the texts. The issue of a multiple theme is closely associated with a textual theme, as the latter is the major contributor to the former. As can be expected, a number of studies found that multiple themes are preferred by less proficient writers (Kim, 2012; Nam & Park, 2015; Park & Lu, 2015; Park & Nam, 2015).

Another salient feature that showed a clear distinction is the marked theme. As noted above, marked choice of a theme occurs for a contextual reason to abandon the normal choice. Hewings (2004) stated that a simple listing of unmarked themes makes a text monotonous. This argument is supported by the studies which report that experienced writers tend to use marked themes more frequently (Hawes & Thomas, 1997; Jalilifar, 2010; Lee, 2008; Park & Lu, 2015; Park & Nam, 2015; Ping, 2007). There exists, however, mixed evidence as to the relationship between learner proficiency and the use of the marked theme. Kim (2012), for example, found that Korean EFL students largely

preferred marked themes compared to native counterparts.¹ The conflicting findings of the previous literature thus demands further study.

The topical theme, an essential ingredient of a thematic structure, has also been at the core of research interests. Analyses on the semantic properties of topical themes first focused on genre comparison. Francis (1990) examined semantic categories of topical themes in different genres of newspapers, and found that 'news reports' preferred more 'concrete things' or 'people' as topical themes than 'editorials' and 'letters of complaint' did. Vande Kopple (1991) and Fries (1995) also showed that a writer's choice of topical theme tends to be affected by genre of writing. In the theme analysis of L2 writings, Lee (2008) found the EFL students' strong preference for personal pronouns, animate or concrete things in the topical position, which resulted in personalizing their essays. Similarly, Kim's (2012) comparative study confirmed the learners' tendency to use more pronouns and concrete themes. The literature has revealed that immature writers' frequent employment of personal contents in the topical position could make it difficult for the formal written texts to accomplish the intended goal of writing (Brown & Yule, 1983; Fries, 1995; Lee, 2008; Thompson, 2004).

2.2.2. Theme analysis in EFL contexts

Although considerable research efforts have been made to assess the effects of thematic choice on the quality of academic writings, theme analysis is an emerging area in L2 writing research. A few attempts focused on the problems that writers in English as a foreign language (EFL) context face in managing the structure of their texts (Herriman, 2011; Mauranen, 1996; Wang, 2007). In terms of comparison across proficiency, Hawes and Thomas (1997) analyzed essays written by Malaysian students attending tertiary education from lower-intermediate to advanced level, and reported the advanced learners' distinguishing characteristics in their thematic choice. For instance, advanced writers used marked themes more frequently and were reluctant to employ concrete nouns as a topical theme. Examples of this type of research also include a couple of analyses undertaken in the Korean EFL context (Lee, 2007, 2008).

Lee's (2007) research on the corpus of Korean college students' argumentative essays focused on the aspect of the textual theme. His study reveals that the EFL essays tend to depend much on the textual theme as well as display register inappropriateness of connectors. The Korean students in his study resorted to connectors having characteristic of the spoken register, which rendered their writings rather informal. This result is, in general, consistent

¹ This unexpected finding might be due to the fact that Kim's (2012) categorization of the theme markedness did not fully reflect Halliday's (1994) theoretical model. He classified all instances of interrogative, exclamatory, and imperative forms into the marked category.

with the findings of most comparative studies (Kim, 2012; Nam & Park, 2015; Park & Lu, 2015; Park & Nam, 2015). With regard to the semantic properties of a topical theme, researchers have highlighted less skillful writers' strong preference for personalized features in formal writing (Fries, 1995; Kim, 2012; Lee, 2008; Thompson, 2004). Lee (2008), for example, found that Korean EFL students are apt to personalize their essays, often employing discourse features of narratives in topical positions.

Recently, Park and Nam (2015) adopted a software program for automatically analyzing the thematic structure. Sample sentences extracted from Korean students' essays and essays collected from an American university were analyzed based on four key features: theme length, theme function, theme plurality and theme role. The result showed that the Korean EFL learners are in general limited in thematic choice. The researchers call for subsequent studies to include comprehensive consideration of learner variables and qualitative analysis in order to thoroughly understand the relationship between the quality of writing and the choice of thematic structure.

As discussed so far, compared to a variety of literature on the native English speakers' thematic choices, relatively less attention has been paid to the EFL learners'. Considering that writing is one of the biggest impediments for non-native learners, more research is necessary to achieve a deeper understanding of how L2 writers employ thematic structure to organize texts. It also needs to be noted that the differential impact of learners' proficiency on thematic outcomes remains elusive. This can be at least in part attributable to the fact that statistical analysis has been beyond the scope of most of the previous research. This methodological challenge is common in the literature on systemic functional approach. As a result, there is insufficient information on the effect of proficiency level, which is vital for the improved L2 writing pedagogy. Currently, several attempts are being made to develop automatic computer tools for analyzing a large amount of data (Park & Lu, 2015; Schwarz et al., 2008). Such automatic analysis, however, would need to be supplemented by detailed examination of individual instances, which is indispensable for uncovering the motives behind the unique patterns. As a way of bridging these gaps, the present study attempts to scrutinize Korean EFL essays from different proficiency levels in comparison with native writers' texts, analyzing an adequate number of texts for performing statistical analyses.

3. METHODOLOGY

3.1. The Texts

The data used in the present study are drawn from the Korean Learner Corpus of English Argumentative Essays (KLCEAE) (Oh & Kang, 2013) and the Louvain Corpus

of Native English Essays (LOCNESS). The KLCEAE is comprised of English argumentative essays written by college students from several universities in Korea. The LOCNESS, which was compiled by the team of Sylviane Granger for reference in the analysis of International Corpus of Learner English (ICLE), includes mainly argumentative essays written by native English speakers.

The KLCEAE consists of a total of 709 essays that are divided into three proficiency levels.² The assessment of the essay is based upon the holistic grading guided by the TOEFL iBT test writing rubric. Three experienced raters, including two English native speakers with an average of 12 years of teaching experience at the tertiary level in the English as a second/foreign language contexts, evaluated all the essays. For the present study 50 essays were extracted each from the three proficiency levels of the KLCEAE as well as from the LOCNESS, totaling 200 essays chosen for analysis. Following the previous studies which considered essay topics as a critical factor in theme analysis (Aijmer, 2002; Herriman, 2011; Hewings, 2004; Hinkel, 1995), the selection of the essay was based on the topics the essays dealt with. The essays covering the same topics in both the corpora (e.g., euthanasia, alcohol problem, abortion and environmental issues) were primarily selected. In order to ensure topic comparability, essays dealing with issues that are specifically more relevant to either of the two writer groups (e.g., Korean political issues or the US government) were excluded. As Biber (1993) argued, the difference in text length can have a negative effect on the corpus analysis; therefore, the native essays over 1000 words were eliminated. Table 1 displays the numbers of essays and words contained in the three levels of the learner corpora and the native corpus used in the present study.

TABLE 1
The Four Corpora of the Study

Number	BL	IL	AL	NS
Essays	50	50	50	50
Words	11291	16405	22577	22581
Words/Essays	225.8	328.1	451.54	451.62
T-units	1052	1141	1397	1290

Note. BL = Beginner level, IL = Intermediate level, AL = Advanced level, NS = Native speakers

3.2. Data Analysis Procedure

The data analysis procedure involved the following steps. First, all the texts were

² The size of the three sub-corpora of the KLCEAE is as follows: the Advanced Level consists of 95 essays with 49,400 words, the Intermediate Level contains 492 essays with 169,441 words, and the Beginner Level is comprised of 122 essays with 27,944 words.

segmented manually into T-units. The t-unit, defined by Hunt (1965) as “an independent clause plus any subordinate clause(s) attached to it” (p. 141), has been adopted by a number of SFG researchers. The second step was to divide each of the t-units into a theme and a rheme. Next, each theme was manually categorized according to the classification of various theme features. The present study adopted four main features of theme based on the previous research (Eggins, 2004; Halliday, 1994; Thompson, 2004) (See Table 2). The study also investigated the specific thematising structures adopted from Thompson (2004) (see Table 3) and the distributions of eight semantic categories of topical themes on the basis of Francis’s (1990) and Lee’s (2008) classifications (see Table 4). In the examples below, themes are underlined, with textual themes italicized and interpersonal themes in boldface.

TABLE 2³
Theme Features

Feature Name	Category	Definition and Example
Theme length		The number of words in the theme
Theme function	Textual	The theme links to the previous sentences (e.g., that is, also, meanwhile)
	Interpersonal	The theme expresses the speaker’s attitude (e.g., probably, usually, fortunately)
	Topical (ideational)	The theme expresses the contents (e.g. the best idea)
Theme plurality	Single	Topical function only (e.g., A slight feeling of fear of his father, mixed with awe, comes through many of his letters.)
	Multiple	A textual and/or an interpersonal theme precedes the topical theme (e.g., Therefore I may as well be cheerful.)
Theme markedness	Unmarked	Normal choice of a theme (e.g., This teapot was what the duke gave to my aunt.)
	Marked	The theme breaks the unmarked norm (e.g., Until the arrival of the remittance I am cut off from my home.)

Based on the categories of various theme types and characteristics shown above, the theme of each t-unit was carefully analyzed and classified with regard to its traits. Following the notational conventions of Halliday and Matthiessen (2013), the results of the thematic analysis were coded in tabular data using Microsoft Excel.

In the next step, for comparisons across corpora, the frequencies of theme features per 1,000 t-units were calculated in each subclass and each semantic category. To assess the

³ This table is adapted from Park and Nam’s (2015) and includes examples taken from Halliday and Matthiessen (2013).

statistical significance of frequency differences between the non-native and the native groups as well as among the different levels of the former, chi-square tests were run by SPSS (Statistical Package for the Social Sciences). Wherever relevant, the actual texts were carefully examined to identify qualitative differences, if any, in the theme usage across the different levels of writing.

TABLE 3
Thematising Structures

Structure	Description and Example
Thematic equatives	The form 'Theme = Rheme' with the '=' expressed by the predicator 'be' (e.g., <u>What you need to</u> do is to write me an official letter asking for an extension.)
Predicated theme	The 'it' acts as a place-holder for the subject of the predicator 'be' in the first clause and the real subject is the second clause. It is possible to re- write the two clauses in the form of a single clause. (e.g., <u>It wasn't until 1986</u> that we finally came back to work in the UK.)
Thematised comment	The comment in the 'it'-clause is not a meaning component of the second clause, therefore, it is impossible to re-write them in the form of a single clause. (e.g., <u>It is true</u> that it took five years to do so.)
Proposed theme	The theme as a separate constituent, and then substituting a pronoun in the appropriate place in the following clause. (e.g., <u>Happiness</u> , that's what life is about.)

TABLE 4
Semantic Categories of Topical Theme

Semantic Category	Example
1st person	I, we
2nd person	you
Animate participants	racehorses, the dogs
Global theme	euthanasia, the cloning of embryos
Abstract concept	the plan, the reason
Anaphoric nominalization	this traumatic event, this hypocritical view
Text reference	that question, this instrument
Spatio-temporal expressions	in March of 1989, in America

In the next step, for comparisons across corpora, the frequencies of theme features per 1,000 t-units were calculated in each subclass and each semantic category. To assess the statistical significance of frequency differences between the non-native and the native groups as well as among the different levels of the former, chi-square tests were run by SPSS (Statistical Package for the Social Sciences). Wherever relevant, the actual texts were carefully examined to identify qualitative differences, if any, in the theme usage across the different levels of writing.

4. RESULTS AND DISCUSSION

4.1. Theme Features

4.1.1. Theme length

Table 5 describes the comparison on theme length among the three levels of Korean EFL learners and the native counterparts.

TABLE 5
Theme Length Across Corpora

	N	Min	Max	<i>M</i>	<i>SD</i>
BL	1052	1	27	2.99	2.69
IL	1141	1	25	3.61	3.04
AL	1397	1	26	4.33	3.83
NS	1290	1	34	4.04	3.98

Note that the mean length of a theme gradually increases with the growth of learner proficiency. The beginner-level learners employ the shortest theme (2.99), and the mean value of theme length grows consistently in the intermediate (3.61) and advanced groups (4.33). One interesting finding is that the mean theme length in the advanced learner writing is even longer than that in the native writing (4.04).⁴

In terms of the standard deviation of theme length, on the other hand, the native corpus marks the highest value (3.98) among the four corpora. This implies that the native writers show the greatest variation in the length of theme and thereby a wider choice in theme selection. This result supports previous studies reporting that the values of theme length and standard deviation in learner data are smaller than those of native corpora (Nam & Park 2015; Park & Lu, 2015; Park & Nam, 2015). Also note that among the three learner groups, the value of standard deviation increases steadily as learner's level of proficiency increases (i.e., 2.69 in BL, 3.04 in IL, and 3.83 in AL). It thus seems that the advanced learners are not just better at producing a longer theme, but also diversifying thematic structures.

A close look at the writing data confirms that the Korean EFL learners, at lower levels in particular, are relatively limited in the aspect of theme length while the native writers show more flexibility. In line with the finding by Nam and Park (2015), longer themes in the native writing typically include the long noun phrases (NPs) with condensed

⁴ It also needs to be noted, though, that the maximum theme length in the AL (26) is much shorter than that in the NS (34), which suggests a greater variation of the latter corpus in theme length.

information. For example, NPs such as *the opportunity to learn about the vast world around us* (NS #08) or *people who inherit money or live off the interest of investors* (NS #21) act as the theme of a clause, providing dense and specific information about the subject. The shorter themes in the native texts often show the use of demonstrative pronouns such as *this* or *that*, which function to connect the previous argument with the following clause. The Korean EFL data, on the other hand, rarely show such long NPs or demonstrative pronouns in the thematic position. Instead, the non-native writers, especially at the beginner and intermediate levels, tend to employ person pronouns (e.g., *I, they*) or personal agents (e.g., *many people, some feminists*) as short themes. Even the advanced learners, who show comparable performance to the native writers in terms of the average theme length and standard deviation, display an interesting difference in the type of the longer themes that they prefer. Instead of long NPs found in the NS data, longer themes in the AL tend to take the form of a dependent clause such as *if you are a businessman and your partner wants to contract with you for profits* (AL #02) and *even though there are class evaluations conducted by government officials* (AL #46).

4.1.2. Theme function

In this subsection, the proportions of three types of theme function, i.e., topical themes, textual themes, and interpersonal themes, are compared among the four corpora. Table 6 presents the raw and the normalized frequency of occurrences of each theme function.

TABLE 6
Density of Theme Functions Across Corpora

	Topical	Textual	Interpersonal
BL	1030 (711.3)	392 (270.7)	26 (17.9)
IL	1137 (715.9)	423 (271.5)	28 (17.9)
AL	1389 (702.2)	538 (271.9)	51 (25.7)
NS	1281 (811.7)	256 (162.2)	41 (25.9)

Note. Numbers in parentheses are frequencies per 1,000 T-units.

The table reveals a clear difference in density of the topical and the textual themes between the Korean EFL learners and the native speakers. Compared to the natives, all of the three learner groups show a considerably more frequent use of the textual theme, while their use of the topical theme is less frequent. What is noteworthy is the fact that the learners' performances in this regard exhibit little change according to the proficiency level. The results of the chi-square tests between the learner groups as a whole and the NS confirm statistically significant differences between the two groups in

the proportion of the topical ($\chi^2 = 121.04, p < .001$) and the textual themes ($\chi^2 = 40.41, p < .001$). The learners' overuse of the textual theme is a finding consistent with previous research (Green et al., 2000; Jalilifar, 2009; Kim, 2012; Lee, 2007; Nam & Park, 2015; Park & Nam, 2015). The learner and the native groups, however, were not significantly different in the use of the interpersonal theme ($\chi^2 = 1.42$). This result is in line with Nam and Park (2015), but conflicts with a couple of other studies reporting that the frequency of interpersonal themes is relatively lower in non-native texts compared to native texts (Jalilifar, 2010; Kim, 2012).

The following excerpts are presented to provide a greater understanding of the Korean EFL learners' excessive dependence on the textual theme as an orienting device. The first example is extracted from an EFL text at the advanced level. (The theme of each T-unit is underlined, and the textual theme, italicized.)

(1) (AL#03)

Korea has a low rate of reported for sex crime. *The reported rate* is only 6%, *and the indictment rate* is lower than 45%. *Furthermore, imprisonment rate* is lower than the indictment rate. *In other words, only 1% of the entire offenders* are sentenced to jail, *and even fewer people's information* is released on the internet. *In other words, there are many hidden criminal that we don't know. For that reason, it is a big mistake to think that the notify law enforcement can preclude crime.*

In the above extract, six successive textual themes (i.e., *and, furthermore, in other words, and, in other words, for that reason*) are employed in the thematic position. As can be seen in this example, however, the marked overuse of textual themes has a negative effect on building coherence of a text, and can be considered as one of the major barriers that deter non-native writers from constructing well-organized texts (Lee, 2007).

Let's compare excerpt (1) with the following written by a native writer:

(2) (NS#31)

Susan B. Anthony originally led the cause because she felt that women should be given equal treatment in society. The ideas behind "feminism" were rooted in the ideal that all men, and women, were created equal and therefore should be treated so in the society in which they chose to live. Men had the exclusive rights to the vote, employment opportunities and other societal features that empower a human being to be innately whole. By this I do not mean that by receiving the right to vote and having a chance to get a good job you have found the meaning of life.

No textual theme is employed in excerpt (2); the text, however, still maintains the logical connection between the clauses, with diverse topic-related topical themes controlling the flow of information. According to Halliday (1994), an unmarked theme also gives continuity to the texts and maintains coherence between sentences. The theme or the rheme of the previous clause can be placed in the thematic position of the following clause and this relation connects the clauses.

The analysis also found substantial differences in the use of textual themes according to the learners' varying levels of proficiency. The result reveals that the beginner group uses the most limited types of connectors such as *and*, *so* or *but*, as displayed in excerpt (3):

(3) (BL#01)

Society make light of lives about to kill their own baby. So I object the abortion. There is exception as I referred. But most of the abortion is just done by their parent's carelessness.

This result is not unexpected considering that these informal connectors have often been considered as an index of low-level learners' 'syntactic immaturity' (Crowhurst, 1987; Reid, 1992; Rutherford, 1987; Scarcella, 1984), and of the spoken discourse rather than the written (Biber, 1986; Chafe, 1984; Granger & Tyson, 1996). In fact, the informal connectors as textual themes are widely found in the Korean EFL data; it may also be related to what Hasselgren (1994) referred to as "lexical teddy bears," which describes the infelicitous usage resulting from the overdependence on the familiar, that is, the expressions learned at the early stage.

Another characteristic of the non-native writers is to adopt ordinal numbers (e.g., *first*, *second*, *third*) as sequential "frame markers" (Hyland, 1998) in the textual theme position. Notably, this feature is the most frequent at the intermediate level, as shown in the following excerpt:

(4) (IL#09)

There are four reasons why I am against. First, suicide is not allowed. Euthanasia is a kind of suicide. Human rights do not include the right to take one's own life. Second, an end cannot justify a means. Even if euthanasia is right, doctor's crime must be punished. Who shall be responsible for it? Euthanasia raises such a problem. Third, there is some possibility of killing treatable patients, because the discernment, whether the disease can be treated or not, is especially difficult. Finally there are worries about a trend to make light of one's life.

In excerpt (4), four reasons for the writer's position on the given issue are listed with sentence-initial ordinal numbers. Noteworthy is that the learners' preference for the ordinal numbers is in clear contrast to the native texts where almost none is found in the thematic position. Frequent use of sentence-initial ordinal numbers in the textual position thus appears to be an indicator of learner writing.

4.1.3 Theme plurality

The current subsection deals with Halliday's (1994) distinction between simple and multiple themes in accordance with their internal structures. Table 7 illustrates the raw and the normalized frequencies of theme plurality across the four corpora.

TABLE 7
Density of Theme Plurality Across Corpora

	Single	Multiple	N/A ⁵
BL	618 (587.4)	412 (391.6)	22 (20.9)
IL	692 (606.4)	445 (390.0)	4 (3.5)
AL	820 (586.9)	569 (407.3)	8 (5.7)
NS	994 (770.5)	287 (222.4)	9 (6.9)

Note. Numbers in parentheses are frequencies per 1,000 T-units.

As shown in the table, both the native and the non-native groups show a more frequent use of the single themes than the multiple themes. This tendency is much stronger, however, in the native group. Note also that the native group's use of multiple theme is noticeably less frequent (222.4) than that of the three non-native groups, who display relatively equivalent frequencies (391.6, 390.0, 407.3). The chi-square test revealed statistically significant differences in the frequencies of the single ($\chi^2 = 129.40, p < .001$) and the multiple themes ($\chi^2 = 127.19, p < .001$) between the two groups. In other words, the Korean EFL learners are more dependent on the multiple themes (and less on the single themes) compared with the native writers.

Structurally, a multiple theme must have at least one or the other of a textual or an interpersonal theme in addition to the topical theme. In this sense, the frequencies of single and multiple themes are closely related to theme functions discussed earlier. As discussed earlier, the learner groups highly overused textual themes compared with the native writers. Considering their preference for the textual function, the learners' heavy

⁵ N/A in Tables 7 and 8 refers to the clauses which could not be analyzed into a theme or a rheme due to the incomplete sentence structure or grammatical errors.

reliance on the multiple theme only seems natural. The results of chi-square tests verify that the distribution of the single themes ($\chi^2 = 130.66, p < .001$) and the multiple themes ($\chi^2 = 128.21, p < .001$) is significantly different among the four groups.

The existing literature has provided mixed evidence on the relationship between the frequencies of multiple theme and learner proficiency. Most studies suggested that learners with lower proficiency overuse multiple themes in comparison to those with higher proficiency (Kim, 2012; Nam & Park, 2015). North's (2005) comparison between UK students majoring in arts and science, by contrast, concluded that students studying arts who used significantly higher numbers of multiple themes received a higher grade in writing assessment. The art majoring students were significantly more likely to use both textual and interpersonal theme, which contributed to making the argument explicit and expressing the writer's viewpoint.

4.1.4 Theme markedness

This subsection reports the results of analysis on the frequency of unmarked and marked themes. Table 8 describes the raw and the normalized frequencies of each theme type.

TABLE 8
Density of Theme Markedness Across Corpora

	Unmarked	Marked	N/A
BL	826 (785.1)	204 (193.9)	22 (20.9)
IL	837 (733.5)	300 (262.9)	4 (3.5)
AL	1023 (732.2)	366 (261.9)	8 (5.7)
NS	979 (758.9)	302 (234.1)	9 (6.9)

Note. Numbers in parentheses are frequencies per 1,000 T-units.

From the data shown in the table, there seems to be no salient difference between the learner and the native groups' selection of theme markedness. The chi-square tests prove that the Korean EFL learners are not significantly different from the native counterpart in the use of the unmarked ($\chi^2 = 0.58$) and the marked ($\chi^2 = 0.35$) themes. This overall correspondence contradicts earlier studies arguing that non-native or novice writers tend to underuse marked themes (Hawes & Thomas, 1997; Jalilifar, 2010; Lee, 2008; Park & Lu, 2015; Park & Nam, 2015; Ping, 2007). What is worth noting, however, is that when the chi-square test was run among the four groups, the differences turned out to be statistically significant. That is, adding learner's proficiency level as a variable,

significant differences were confirmed regarding the density of the unmarked ($\chi^2 = 11.47$, $p < .005$) and the marked themes ($\chi^2 = 19.48$, $p < .001$). In this respect, the outcomes of the present study support the conventional wisdom that learners' choice of theme markedness is influenced by their level of proficiency.

4.2. Thematising Structures

The comparisons of the frequency of the three thematising structures in the native and the students' writing are described in this section. Table 9 shows the raw and the normalized frequencies of each thematising structure.

TABLE 9
Density of Thematising Structures Across Corpora

	Proposed Theme	Predicated Theme	Thematic Equative
BL	9 (8.5)	14 (13.3)	17 (16.1)
IL	4 (3.5)	34 (29.7)	11 (9.6)
AL	1 (0.7)	49 (35.0)	19 (13.6)
NS	11 (8.5)	32 (24.8)	39 (30.2)

Note. Numbers in parentheses are frequencies per 1,000 T-units.

The result of the chi-square tests indicates that the L1 and L2 writers are considerably different in their usage of thematic equatives ($\chi^2 = 16.11$, $p < .001$) and proposed themes ($\chi^2 = 3.99$, $p < .05$) whereas the difference is not significant with predicated themes ($\chi^2 = .18$). When the chi-square tests were carried out to evaluate the differences among the four groups, on the other hand, the results point to significant differences in all the three thematising structures: proposed themes ($\chi^2 = 11.27$, $p < .05$), predicated themes ($\chi^2 = 11.73$, $p < .005$) and thematic equatives ($\chi^2 = 17.48$, $p < .001$). In contrast to the non-significant difference between the native and the non-native groups, the four groups were significantly different in the density of predicated themes (as well as the other two structures). This may be attributed to the beginner group's much lower use of the predicated themes, as seen in the raw frequency.

The relative underuse of these specialized structures in the EFL corpus needs to be considered in relation to the functions they perform. As described in Table 3, the thematic equative takes the form of 'Theme = Rheme', with the '=' articulated by the predicator 'be' (e.g., *what Einstien did to humanity's philosophy was change the notion that closure could be obtained*, NS#12) (Thompson, 2004, p. 126). In a thematic equative, the starting point is usually a question that the writer or speaker assumes the

audience may want to ask at that stage of the text. By completing his or her own message by answering the question, the thematic equative “makes explicit the interactive consideration of the audience” (Thompson, 2004, p. 127). Predicated themes allow the writer to pick out a single constituent and give it emphatic thematic status (e.g., *it is only in very recent times that genetic manipulation has been seen to have bad as well as good sides*, NS#46). Similar to thematic equatives, predicated themes guide the audience to pay attention to a particular constituent by selecting a specific element and emphasizing it. In this vein, it can be said that native writers are better able to interact with the readers by using specific thematising structures.

4.3. Semantic Properties of Topical Theme

This section analyzes the semantic properties of topical themes in the four corpora. Table 10 describes the raw and the normalized frequencies of eight different types of semantic properties in each corpus.

TABLE 10
Density of Semantic Properties of Theme Across Corpora

	BL	IL	AL	NS
1 st person	270 (256.7)	189 (165.6)	146 (104.5)	127 (98.5)
2 nd person	8 (7.6)	2 (1.8)	9 (6.4)	9 (7.0)
Animate participants	0 (0.0)	0 (0.0)	9 (6.4)	5 (3.8)
Global theme	50 (47.5)	62 (54.3)	68 (48.6)	56 (43.4)
Abstract concept	176 (167.3)	161 (141.1)	298 (213.3)	299 (231.8)
Anaphoric nominalization	3 (2.9)	4 (3.5)	11 (7.9)	10 (7.8)
Text reference	127 (120.7)	170 (148.9)	173 (123.8)	157 (121.7)
Spatio temporal	33 (31.4)	32 (28.0)	31 (22.1)	33 (25.5)

Note. Numbers in parentheses are frequencies per 1,000 T-units.

According to the chi-square test results, the first person pronouns ($\chi^2 = 36.96, p < .001$) are significantly overused, and abstract concepts are significantly underused ($\chi^2 = 18.48, p < .001$) in the Korean EFL texts compared to the native texts. Considering the comparable essay topics of the two corpora, this finding is especially noteworthy. In fact, this result supports the previous studies showing that immature writers frequently use the first/second person pronouns as a theme (Brown & Yule, 1983; Kim, 2012; Lee, 2008; Thompson, 2004) and less prefer abstract concepts in the thematic position (Lee, 2008).

The chi-square tests comparing the four corpora also indicated significant differences with regard to the first person pronouns ($\chi^2 = 145.60, p < .001$) and the abstract concept ($\chi^2 = 40.52, p < .05$). Additionally, there was a statistically significant difference in the density of animate participants ($\chi^2 = 13.00, p < .005$). The literature suggested that the recurrent use of the first and second person pronouns in theme position represents a key characteristic of speech rather than writing (Fries, 1995; Thompson, 2004) while the employment of abstract concepts including anaphoric nominalization is in general considered characteristic of the formal written discourse (Lee, 2008). In this respect, the current result seems to reflect the lower level learners' insufficient knowledge on the register and genre.

In the following, typical examples from the each level of the EFL texts are shown for comparison purpose. At the beginner and the intermediate levels, the first person pronouns repeatedly occur as topical themes as shown in excerpts (5) and (6). (The first person pronouns placed in the thematic position are in boldface.)

(5) (BL#13)

I think that the teacher-assessment system is very good idea. (...) **We** have to evaluate person by major group. **I** heard that many teachers are against the teacher-assessment system. **I** cannot understand their insistence. If they are evaluated, they could know what to change and how to teach well.

(6) (IL#23)

I think writing assessment is not necessary for college admission. **I** think it cannot measure student's all ability that college requires. It is true writing assessment can distinguish students.

In the excerpts above, first person pronouns are successively placed in the thematic position, presenting the writer's personal experiences as well as opinions on the given issue. As in both of the examples, the student writers' use of the first person pronouns often involves the sentence-initial epistemic expression *I think*, which renders the arguments subjective and personal. This type of recurrent direct involvement of the writer as a text participant seems to reflect their insufficient knowledge about formal writing conventions.

In the advanced texts, in contrast, the first person pronouns are relatively infrequent. In excerpt (7) below, for example, an advanced EFL writer develops his/her argument without employing first person pronouns as a theme.

(7) (AL#44)

Generally, people do not hesitate to say that the writing assessment can show students' academic abilities and potentialities, yet the two apparently contradictory opinions are never in agreement. However, considering the positive effects of the writing assessment, it should be allowed as a means of college admission test.

In (7), the themes are related to people in general (people) or an abstract concept (*the two apparently contradictory opinion*). The example shows how advanced writers make their claims sound rather objective by engaging the majority and abstractions in the thematic position of the arguments. The relatively uncommon use of the first person pronoun in the high-performing group may thus function to increase persuasiveness of their argument. In this respect, they seem to be more familiarized with the academic writing convention regarding the use of the personal pronouns.⁶

As illustrated in the examples above, the present study revealed remarkable frequency variations across the three levels of learner groups, with the frequency of the first person pronouns decreasing as the learner's proficiency develops. This suggests that dependence on the personal pronouns is reduced over time as a result of improved English proficiency.

5. CONCLUSION

The present study explored the thematic structure in Korean EFL learners' writing compared to native texts, with particular focus on learners' proficiency as a variable. The research aimed to find out whether the learners' improvement in overall proficiency is accompanied by more diverse and appropriate use of the thematic structure. For this purpose, the study compared the frequencies of various theme features in three different levels of the learner as well as the native writing, and whenever necessary closely examined the actual writing data.

The findings of the present study clearly show that textual themes, multiple themes, and the first person pronouns as topical themes are highly overused in the Korean EFL texts compared to those of the native English writers. This result is in accord with the previous literature that explored the thematic structures of English texts written by non-native learners (Granger & Tyson, 1996; Green et al., 2000; Jalilifar, 2009; Lee, 2007; Lee, 2008;

⁶ This is not to say, however, that the academic writing convention prohibits the use of the first/second person pronouns. In fact, some researchers (e.g., Chang & Swales, 1999; Hyland, 2004) have pointed out the need to expose the writer's explicit presence in formal essay writing, which can add the 'interactional aspects' (Brown & Yule, 1983) to the discourse. Using the first/second person pronouns in academic writing may thus have its advantages although the writer's direct involvement in discourse is generally uncommon in this register.

Park & Nam, 2015). The result also indicates that special thematising structures such as thematic equatives and preposed themes and abstract concepts as topical themes are underused in the EFL writing. What the current study contributes most to this line of research is the discovery that proficiency is indeed a significant factor which affects the performances of learners in various respects. First of all, learners become more flexible in length of theme as their proficiency develops. Secondly, learners at different levels exhibited significant differences in the proportion of diverse theme features: textual themes, topical themes, theme plurality, and theme markedness. In addition, the thematising structures and the semantic properties including the personal pronouns, animate participants, and abstract concepts, were used differently by the four writer groups. In sum, the comparison generally revealed that the non-native learners' thematic choice changes as their proficiency develops. Nevertheless, the proficiency development does not necessarily involve the learners' proper use of textual or multiple themes. While learners at the beginner level were the most limited in their choice of textual connectors, the tendency to overuse them remained the same even at the advanced level. By contrast, the overuse of the first person pronouns as topical themes noticeably decreases as proficiency grows, presumably along with the enhanced understanding of the discourse practices of the formal writing register.

From a pedagogical point of view, the present study provides some useful implications for the Korean EFL writing context. First of all, learners' awareness needs to be raised of the importance of the theme selection and its role in constructing logical and coherent text. For the next step, explicit teaching efforts along the lines suggested by some researchers would be useful, which focus on eliciting the appropriate use of thematic structures (Lee, 2008; Park & Nam, 2015) and developing the ability to manipulate thematic contents based upon the understanding of conventional generic pattern (Gosden, 1992). In particular, considering the present finding that the learners' proficiency interacts with their thematic choices, the educational efforts should take into account the level of proficiency to provide effective help. The instruction may aim at helping beginners to develop consciousness of the critical role of the initial position of a clause, and then become familiar with a wide variety of theme types they can choose from as well as the rhetorical consequences of their choice. Given that several aspects of thematic structures are still problematic for advanced learners (e.g., textual themes, thematic equatives and preposed themes), it seems also necessary to provide higher-level learners with explicit instruction specifically targeting those particular theme features. Such teaching efforts that are tailored for individual learners' needs are expected to assist them to equip themselves with a successful linguistic strategy for enhancing textual cohesion and coherence through thematic choice, thereby gaining greater control of the writing process and boosting confidence in L2 writing.

While the present study adds to the deeper understanding of L2 writers' thematic organization, it is not without limitations. There might be individual differences, for example, in the thematic choice of both non-native and native writers, but such individual differences have not been considered in the current analysis. Further research that focuses on individual variables such as stylistic preference or writing habits in the choice of themes could help provide a more varied and dynamic picture of thematic organization in writing.

REFERENCES

- Aijmer, K. (2002). Modality in advanced Swedish learners' written interlanguage. In S. Granger, J. Hung & S. Petch-Tyson (Eds.), *Computer learner corpora, second language acquisition and foreign language teaching* (pp. 55-76), Amsterdam: John Benjamins.
- Biber, D. (1986). On the investigation of spoken/written differences. *Studia Linguistica*, 40, 1-38.
- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8(4), 243-257.
- Bloor, M., & Bloor, T. (1992). Given and new information in the thematic organization of text: An application to the teaching of academic writing. *Occasional Papers in Systemic Linguistics*, 6(1), 33-43.
- Bolton, K., Nelson, G., & Hung, J. (2002). A corpus-based study of connectors in student writing: Research from international corpus of English in Hong Kong (ICE-HK). *International Journal of Corpus Linguistics*, 7(2), 165-182.
- Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.
- Chafe, W. (1984). How people use adverbial clauses. *Berkeley Linguistics Society*, 10, 437-449.
- Chang, Y. Y., & Swales, J. (1999). Informal elements in English academic writing: Threats or opportunities for advanced non-native speakers? In C. Candlin & K. Hyland (Eds.), *Writing: Texts, processes and practices* (pp.145-167). New York: Longman.
- Crowhurst, M. (1987). Cohesion in argument and narration at three grade levels. *Research in the Teaching of English*, 21(2), 185-201.
- Eggins, S. (2004). *Introduction to systemic functional linguistics* (2nd ed.). New York: Continuum.
- Fontaine, L., & Kodratoff, Y. (2011). The role of thematic and concept texture in scientific text: Comparing native and non-native writers of English. *La Revue du GERAS Journal*, 5, 1-23.
- Francis, G. (1990). Theme in the daily press. *Occasional Papers in Systemic Linguistics*, 4, 51-87.

- Fries, P. H. (1983). On the status of theme in English: Arguments from discourse. *Forum Linguisticum*, 6(1), 1-38.
- Fries, P. H. (1995). Theme, methods of development, and texts. In R. Hasan & P. H. Fries, (Eds.), *On subject and theme: From the perspective of functions in discourse* (pp. 317-359). Amsterdam: John Benjamins.
- Ghadessy, M. (1999). Thematic organisation in academic article abstracts. *Estudios Ingleses de la Universidad Complutense*, 7, 141-161.
- Ghadessy, M., & Gao, Y. (2000). Thematic organization in parallel texts: Same and different methods of development. *Text-Interdisciplinary Journal for the Study of Discourse*, 20(4), 461-488.
- Gómez, M. (1994). Theme and textual organization. *Journal of Words*, 45(3), 293-305.
- Gosden, H. (1992). Discourse functions of marked theme in scientific research articles. *English for Specific Purposes*, 11(3), 207-224.
- Granger, S., & Tyson, S. (1996). Connector usage in the English essay writing of native and non-native EFL speakers of English. *World Englishes*, 15(1), 17-27.
- Green, C. F., Christopher, E. R., & Mei, J. L. K. (2000). The incidence and effects on coherence of marked theme in interlanguage texts: A Corpus-based Enquiry. *English for Specific Purposes*, 19(2), 99-113.
- Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). London: Edward Arnold.
- Halliday, M. A. K., & Matthiessen, C. (2013). *Halliday's introduction to functional grammar* (4th ed.). London: Taylor & Francis.
- Hasselgren, A. (1994). Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary. *International Journal of Applied Linguistics*, 4(2), 237-258.
- Hawes, T., & Thomas, S. (1997). Problems of thematisation in student writing. *RELC Journal*, 28(2), 35-55.
- Herriman, J. (2011). Themes and theme progression in Swedish advanced learners' writing in English. *Nordic Journal of English Studies*, 10(1), 1-28.
- Hewings, A. (2004). Developing discipline-specific writing: an analysis of undergraduate geography essays. In L. Ravelli & R. Ellis (Eds.), *Analysing academic writing: Contextualized frameworks*, London: Continuum.
- Hinkel, E. (1995). The use of model verbs as a reflection of cultural values. *TESOL Quarterly*, 29(2), 325-343.
- Hunt, K. W. (1965). *Grammatical structures written at three grade levels* (NCTE research report no. 3). Urbana, IL: National Council of Teachers of English.
- Hyland, K. (1998). Persuasion and context: The pragmatics of academic metadiscourse. *Journal of Pragmatics*, 30(4), 437-455.

- Hyland, K. (2004). *Disciplinary discourses: Social interactions in academic writing* (2nd ed.). Ann Arbor: University of Michigan Press.
- Jalilifar, A. (2009). Thematic development in English and translated academic texts. *Journal of Language & Translation*, 10(1), 81-111.
- Jalilifar, A. (2010). The status of theme in applied linguistics articles. *Asian ESP Journal*, 6(2), 7-39.
- Kim, J. H. (2012). *Korean advanced learners' theme choices as linguistic option in L2 writing: A comparative study with the native corpus*. Unpublished Master's thesis, Hankuk University of Foreign Studies, Seoul.
- Lee, S. W. (2007). Connectors as textual orienting theme in Korean and Australian student essay writing. *English Teaching*, 62(1), 121-147.
- Lee, S. W. (2008). Investigating ideational thematic content in Korean university student essay writing. *English Teaching*, 63(2), 71-94.
- Lu, X. (2002). Discourse and ideology: The Taiwan issue in the Chinese and American media. In C. N. Candlin (Ed.), *Research and practice in professional discourse* (pp. 589-608). Hong Kong: City University of Hong Kong Press.
- Martínez, I. A. (2003). Aspects of theme in the method and discussion sections of biology journal articles in English. *Journal of English for Academic Purposes*, 2(2), 103-123.
- Mauranen, A. (1996). Discourse competence—evidence from thematic development in native and non-native texts. In E. Ventola & A. Mauranen (Eds.), *Academic writing: Intercultural and textual issues* (pp. 195–230). Amsterdam: John Benjamins.
- Nam, D., & Park, K. (2015). Corpus-based analysis of thematic structure in L2 writing. *Multimedia-Assisted Language Learning*, 18(4), 99-120.
- North, S. (2005). Disciplinary variation in the use of theme in undergraduate essays. *Applied Linguistics*, 26(3), 431-452.
- Oh, S. Y., & Kang, S. J. (2013). The effect of English proficiency on Korean undergraduates' expression of epistemic modality in English argumentative writing. *The Journal of Asia TEFL*, 10(4), 97-132.
- Park, K., & Lu, X. (2015). Automatic analysis of thematic structure in written English. *International Journal of Corpus Linguistics*, 20(1), 81-101.
- Park, K., & Nam, D. (2015). Analysis of thematic structure in L2 writing: A systemic functional perspective. *The SNU Journal of Education Research*, 24, 65-88.
- Ping, A. L. (2007). Developing the message: Thematic progression and student writing. *The Journal of Asia TEFL*, 4(3), 93-127.
- Reid, J. (1992). A computer text analysis of four cohesion devices in English discourse by native and nonnative writers. *Journal of Second Language Writing*, 1(2), 79-107.
- Rutherford, W. E. (1987). *Second language grammar: Learning and teaching*. London: Routledge.

- Scarcella, R. C. (1984). How writers orient their readers in expository essays: A comparative study of native and non-native English writers. *TESOL Quarterly*, 18(4), 671-688.
- Schwarz, L., Bartsch, S., Eckart, R., & Teich, E. (2008). Exploring automatic theme identification: A rule-based approach. In A. Storrer, A. Geyken, A. Siebert & K.-M. Würzner (Eds.), *Text resources and lexical knowledge: Selected papers from the 9th conference on natural language processing* (pp. 15-26). Berlin: Mouton de Gruyter.
- Thompson, G. (2004). *Introducing functional grammar* (2nd ed.). London: Edward Arnold.
- Vande Kopple, W. J. (1991). Theme, thematic progressions, and some implications for understanding discourse. *Written Communication*, 8(3), 311-347.
- Wang, I. (2007). Theme and rheme in the thematic organization of texts: Implications for teaching academic writing. *Asian EFL Journal*, 9(1), 164-176.

Applicable levels: Tertiary

Yerahm Choung
Department of English Education
College of Education, Seoul National University
1 Gwanak-ro, Gwanak-gu
Seoul 08826, Korea
Phone: 02-880-7670
Email: freezism@snu.ac.kr

Sun-Young Oh
Department of English Education
College of Education, Seoul National University
1 Gwanak-ro, Gwanak-gu
Seoul 08826, Korea
Phone: 02-880-7675
Email: sunoh@snu.ac.kr

Received on June 1, 2017
Reviewed on July 15, 2017
Revised version received on August 15, 2017