Stance in the Introductory *it* Construction: A Comparative Study of Argumentative Writing by Korean EFL and English L1 Students

Choongil Yoon†

Dongguk University(Gyeongju Campus)

ABSTRACT

The construction of an introductory *it* followed by an extraposed subject has been found to be frequently used in academic prose. Recent research has focused particularly on rhetorical motivations for the construction and demonstrated that it provides the writer with a means of marking authorial stance while concealing its source. This study investigated how Korean EFL university students used this rhetorical device in their argumentative writing to encode stance, in comparison with a group of English L1 students. Results showed that while the Korean EFL writers used the construction far more frequently to mark attitudinal stance than their native speaker counterparts, its use was more limited in terms of lexical choice and the rhetorical function of depersonalized stance marking. Based on these findings, this paper offers suggestions on how to help EFL writers acquire the multi-faceted usage of the construction.

Keywords: introductory *it*, extraposition, stance marking, Korean EFL writers, academic writing

1. Introduction

Despite the common perception that academic texts should be presented in a faceless and impersonal manner, research has demonstrated that academic writing is in fact full of personal stance representing the author's opinions and attitudes (e.g., Hunston, 2004; Charles, 2007; Hyland, 2005). Appropriate presentation of authorial stance is considered a feature of advanced academic texts as it serves to provide an impression of the author and control the dialogic space in relation to his or her arguments and readers (Hyland, 2016; Jiang, 2015; Martin & White, 2005). At the same time, research has also revealed that there are some specific ways of expressing stance that are more preferred in academic writing than in other registers such as

† Corresponding author: chongal2@hotmail.com

casual conversation (e.g., Biber, Johansson, Leech, Conrad, Finegan, & Quirk, 1999; Gray & Biber, 2012; Hyland & Tse, 2005). For example, in academic texts, certain lexico-grammatical features are often employed to mark subjective stance while maintaining the appearance of objectivity and accuracy.

One such lexico-grammatical feature widely used is a sentence structure consisting of a matrix clause with the pronoun *it* as its subject and an extraposed nominal clause as seen in examples (1) and (2):

(1) *It is highly likely that* nothing significant will be found.
(2) *It is important to* identify hazards systematically.

Extraposition here refers to the process of placing usually a long and heavy nominal clause (either finite or non-finite) at the right of the matrix predicate and inserting the introductory pronoun *it* in its place. The predicate of the main clause (e.g., *highly likely* and *important* in the examples above) often expresses the writer's stance or viewpoint on the propositional content of the extraposed clause that follows. It has been suggested that by using the impersonal pronoun *it* as the subject, this construction (hereafter referred to as introductory *it* construction) allows the author to depersonalize his or her stance and present it as an objective and uncontested view, enhancing the acceptability or validity of the argument being advanced (Herriman, 2000; Hewings & Hewings, 2002; Hyland & Tse, 2005).

However, the appropriate use of the construction in academic text can be challenging—in terms of both the form itself and its stance-marking function—to English as a second or foreign language (L2) writers. As many languages do not have a corresponding grammatical structure to the introductory *it* construction (Jacob, 1995), it may require repeated exposure and use for L2 writers to become familiarized with the lexical and syntactic restrictions in its usage. Even once they gain some lexico-grammatical control in the use of the introductory *it* construction, L2 writers may still experience some confusion about whether and to what extent personal stance can be expressed with the construction (Hewings & Hewings, 2002; Römer, 2009).

Despite the importance of and potential difficulties with the introductory *it* construction in academic writing, relatively little research has been conducted on stance marking with the use of the construction by English as a foreign language (EFL) academic writers, especially Korean L1 writers of English. This study examined a corpus of argumentative essays written by Korean EFL university students to

investigate how the novice academic writers encoded their opinions and affective meanings using the introductory *it* construction in comparison with the use of the same construction observed in a corresponding native speaker (NS) corpus.

2. Introductory *it* Construction in Academic Writing

2.1. Research foci of previous studies

While different structures of an introductory *it* followed by an extraposed clausal subject have received a great deal of research attention from a number of areas of linguistics, there have traditionally been two major research foci on the topic, which both involve a choice between different options. First, many researchers have suggested that the use of the introductory *it* construction is motivated by thematic choice and information packaging involving end-focus and end-weight principles (e.g., Biber et al., 1999; Herriman, 2000; Kaltenböck, 2000; Quirk, Greenbaum, Leech, & Svartvik, 1985). That is, the choice of an extraposed structure (e.g., *it is important to take preventive measures*) over a non-extraposed variant (e.g., *To take preventive measures is important*) is guided by the tendencies in English to place "new" information after "given" information and to shift long and heavy elements toward the end of the sentence. Second, research has demonstrated that there is some correlative interaction between the semantic meaning of the matrix predicate and the clausal type of extraposed subject that occurs with it (i.e., *that*-clause vs. infinitival clause) (Collins, 1994; Herriman, 2000; Quirk et al., 1985; Zhang, 2015).

In English for academic purposes (EAP) contexts, along with the two foci noted above, another aspect has been particularly foregrounded in research: rhetorical motivations for the use of the introductory *it* construction in academic discourse. A number of studies have suggested that the construction is one of the major linguistic features for stance marking, employed more widely in academic writing than in other genres and registers (Biber et al., 1999; Larsson, 2017; Zhang, 2015). The often suggested reason is that while foregrounding the writer's stance by presenting it as the theme of the sentence, the construction hides the human source of (or depersonalizes) the stance expressed by using the impersonal pronoun *it*, thereby giving it the appearance of objectivity and generality (Herriman, 2000; Hyland & Tse, 2005; Zhang, 2015) as illustrated in the example below.

(3) *It may be argued that* the current problems are temporary.

Although the opinion expressed in the matrix predicate is attributable to the writer, the agent of the passive verb in Example (3) is not provided, thereby concealing whose argument is being advanced at least on the surface. As such, the introductory *it* construction provides academic writers with a means of encoding their subjective stance "objectively," thereby making the argument or statement in question less open to rejection or negotiation (Hyland & Tse, 2005).

A considerable amount of research on the construction has been conducted in EAP contexts, although much of it investigated only some subpatterns of the construction (e.g., *it* + predicate + *that*-clauses, or *it* + copula + ADJ) or examined the construction as a secondary focus. Among others, the studies have identified variables that influence the use and frequency of the construction, and stance expressed therein, such as discipline (e.g., Hyland & Tse, 2005; Peacock, 2011), text genre (e.g., Biber et al., 1999; Zhang, 2015) and NS status of the writer (Hewings & Hewings, 2002; Larsson, 2017). Research findings on L2 writers' stance marking in the introductory *it* construction, which are directly relevant to this study, are reviewed in more detail in the next section.

2.2. Stance expressed in the introductory *it* construction

The construct of stance has been examined from various perspectives such as *evaluation* (e.g., Hunston & Thompson, 2000), *appraisal* (e.g., Martin & White, 2005), and *metadiscourse* (e.g., Hyland, 2005) that overlap to various degrees in their roles and linguistic realizations in text (see Biber, 2006; Hyland, 2016; Gray & Biber, 2012 for a quick review of relevant studies of stance). However, in EAP the term *stance* is specifically linked to "the ways in which speakers and writers encode opinions and assessments in the language they produce" (Gray & Biber, 2012, p. 15). In academic writing, stance is considered to play an integral role in engaging readers and persuading them to accept the arguments being presented (Jiang & Hyland, 2018) because the persuasive force of an argument comes from the author's ability to evaluate and analyze the given knowledge as well as her ability to make a propositional claim and provide supporting grounds for it (Jiang, 2015; Wingate, 2012).

Stance in academic prose has been classified into either broad semantic categories or the functional roles of linguistic features in marking stance. For example, while

Biber et al. (1999) break stance into two broad semantic categories of epistemic stance (expressing certainty and likelihood) and attitudinal stance (attitudes, evaluations, personal feelings and emotions), Hyland divides stance into more functional categories of hedges, boosters, attitude markers, and self-mention (see Hyland, 2016 for definitions and examples of each category).

When it comes to stance in the introductory *it* construction, Herriman (2000) offers a useful analytic framework. Examining the uses of the construction in the multi-genre Lancaster-Oslo/Bergen corpus, she classified the matrix predicates occurring in the construction into semantic categories representing epistemic, deontic, dynamic, and evaluative modalities. The study largely confirmed the findings of the previous studies (Biber et al., 1999; Collins, 1994) commonly demonstrating that the type of extraposed clause (e.g., *to*-infinitive or *that*-clause) is to a large extent associated with the semantic meanings of the matrix predicate. For example, while the matrix predicates with epistemic meanings occurred with finite clauses most frequently (e.g., *it is possible that* ⋯, *it is obvious that* ⋯), those expressing dynamic modalities were used frequently with non-finite clauses (e.g., *it is difficult to* ⋯, *it is easy to* ⋯). The researchers attributed the contrast to the factivity and generality of the content represented by each type of clause.

There have been studies that examined the uses of the construction (often along with other lexico-grammatical patterns) by English L2 learners, often in comparison with those by NS writers (student and/or expert writers) (e.g., Ädel, 2014; Ädel & Erman, 2012; Hasselgård, 2009). However, only a small number of studies have focused on the ways in which English L2 students express stance using the construction. Hewings and Hewings (2002) was one of the first studies that investigated L2 students' stance marking with the introductory *it* construction as the main research focus. They compared instances of the construction used in English L2 MBA students' dissertations with those in published journal articles, using a functional classification in line with Hyland's (2016) discussed above. The study found that the L2 students used the construction more often for emphasis and attitude marking, but less for hedging, suggesting the student writers made greater efforts to make their propositions and claims more forceful and persuasive than the expert writers. These findings have largely been confirmed in later studies. In her investigation of the *it is* ADJ + extraposed clause patterns in English essays of writers with different academic writing proficiencies and NS status (German learners of English, NS students and published writers), Römer (2009) also found that NNS students had a greater tendency than NS writers to use "extreme" adjectives such

as *amazing* and *stupid*, expressing strong emotions. In a study comparing Swedish L1 writers' use of the introductory *it* construction with that of NS counterparts, Larsson (2017) also found that the NNS students underused the pattern for hedging. At the same time, however, these studies commonly suggest that factors such as language proficiency, expertise in academic writing, and L1 transfer may play significant roles in influencing the frequency and use of the construction.

2.3. Research questions

Despite its frequent use in academic writing, studies on stance marking with the introductory *it* construction have largely been confined to NS writers and NNS students at advanced levels, often L1 users of European languages which have grammatical structures corresponding to *it*-clauses (e.g., Ädel, 2014; Ädel & Erman, 2012; Hasselgård, 2009; Römer, 2009). There has been relatively little research specifically looking into novice EFL academic writers, including Korean EFL writers, who have no similar grammatical structures in their native language.

Motivated by the lack of research attention, this study examined, both quantitatively and qualitatively, the use of the introductory *it* construction and how stance is encoded in the construction from a small corpus of argumentative writing by Korean L1 university students and compared it with data from a corresponding NS corpus of argumentative writing. The study was guided by the following specific research questions:

1. How frequently are the introductory *it* construction and its structural subtypes used in argumentative writing by the Korean and NS university students?
2. What types of stance are expressed with the construction?
3. How do the two groups of student writers match the stance meanings of the matrix predicate with the clausal types of the extraposed subject?
4. What are, if any, the main patterns of unconventional or erroneous uses of the construction by the Korean university students?

3. Method

3.1. Corpora

For the present study, the argumentative essay component of Neungyule Interlanguage Corpus of Korean Learners of English (NICKLE)[1] was used. NICKLE is a million-word multi-genre corpus compiled as part of the Neungyule-Longman English-Korean Dictionary project in 2009. The source text data were essays written by the first- and second-year undergraduate students at intermediate levels as a course assignment or class activity in multiple universities across South Korea. A total of 286 essays were selected from the corpus based on the essay title and genre provided in the meta-data. For a NS reference corpus, the Louvain Corpus of Native English Essays (LOCNESS)[2] was used, which is made up of essays written by British and American university students who were native speakers of English. Only a part (184 essays) of its argumentative essay component was randomly selected for comparability with the NICKLE subcorpus in terms of genre and size. The vast majority of the essays in both subcorpora were written on popular argumentative topics such as euthanasia, the death penalty, and environmental protection with some unique topics to each corpus (e.g., English learning for NICKLE and European Union for LOCNESS). Table 1 below summarizes the details of the two subcorpora of argumentative essays (hereafter referred to as NICKLE and LOCNESS respectively).

Table 1. Corpus profile

Corpus	Word tokens	Number of essays
NICKLE	148,127	286
LOCNESS	147,777	184

3.2. Procedure

Data retrieval. To extract all instances of the introductory *it* construction, the two corpora were first tagged for part of speech using the Multidimensional Analysis

---

1) The corpus is freely available upon request. For more information about the corpus, go to https://uclouvain.be/en/research-institutes/ilc/cecl/learner-corpora-around-theworld.html.
2) Go to https://uclouvain.be/en/research-institutes/ilc/cecl/locness.html for detailed information on the corpus.

Tagger (Version 1.2) (Nini, 2014). Next, through regular expression queries on AntCont (Version 3.4.3), instances of the pronoun *it* followed by a verb (including modal verbs) were retrieved. All these instances were then manually examined with their broader contexts to select only the valid instances of the construction as defined earlier: introductory *it* + predicate + a finite (*that-* and *wh*-clause) or non-finite (*to*-infinitive and gerundial clause) clause as the extraposed subject. However, syntactically similar structures were excluded such as *it*-clefts (e.g., *it was in Berlin that I first met her*), referring *it* + relative clause (e.g., *it wasn't the book that I'd looked for.*) and referring *it* + adverb infinitive (e.g., *it was excluded to avoid redundancy*). Finally, it should be noted that as one of the research questions concerned unconventional or erroneous uses of the construction, lexically and/or grammatically infelicitous instances were not excluded from analysis as long as they were judged to be intended uses of the introductory *it* + extraposed subject pattern (e.g., *it is sure that one's life is more valuable than money*).

Stance classification. The extracted valid instances were classified according to the types of stance expressed in their matrix predicates and the ways in which stance was presented. For comparability with previous studies of stance, a coding scheme was developed by integrating classifications adopted by Biber (2006), Herriman (2000), Hewings & Hewings (2002), and Hyland (2016). Stance expressed in the matrix predicates were first divided into two broad semantic categories: *Epistemic* and *Attitudinal*. *Epistemic* stance refers to the status of information in a proposition, indicating the writer's opinion of the truth value or factuality of the content of the extraposed clause. It breaks further down into *Likelihood* and *Certainty*. On the other hand, *Attitudinal* stance, indicating the writer's assessment of or personal attitudes toward a proposition, are further divided into *Deontic*, *Dynamic* and *Evaluative* stance. The predicates judged not to fit in any of these semantic categories or not to indicate any particular stance of the writer were all coded as *Others*. Table 2 below presents the definition and examples of each category discussed above.

Next, the matrix predicates were also functionally coded as *Hedged* or *Emphatic*. The predicates were coded as *Hedged* when they came with words or phrases that withhold the writer's full commitment to or tone down her assertion whereas they were coded as *Emphatic* when used with boosters that emphasize the force of the writer's assertion. Examples of each category are also presented in Table 2 below.

Given the classification presented in Table 2 below, a further explanation seems in order for two potential areas of confusion. First, the functional classification of *Hedged* and *Emphatic* may appear to overlap to a large extent with the *Epistemic*

Table 2. Semantic and functional classifications of the introductory it construction

Categories for matrix predicates			Descriptions and examples
Semantic	**Epistemic**		Indicates the writer's opinion of the status (truth value, factivity) of the content of the extraposed clause
	Likelihood		*it is likely/conceivable; it seems; it is doubtful; it is assumed/argued/claimed/suggested*
	Certainty		*it is true/obvious/evident; it is my firm belief; it follows; it is proven/established*
	Attitudinal		Indicates personal attitudes toward or evaluation of the content
	Deontic		Obligation and volition
		Obligation	*it is necessary/essential/mandatory*
		Volition	*it is desirable/preferable; it is my wish/aim/intention; it is recommended*
	Dynamic		Ability or power to carry out a course of action
		Potentiality	Potential for success/failure, the ease or difficulty involved: *it is difficult/impossible/easy*
		Circumstances	Various circumstances connected to the potential success (speed, manner, cost etc.): *it is safe/cheap, it takes a long time*
	Evaluative		Various value judgements
		General evaluation	Opinion about the (un)favorability of the content: *it is good/fortunate, it is a shame/a pity*
		Appropriate-ness	Opinion about the correctness/suitability of the content: *it is reasonable/appropriate/natural*
		Significance	Opinion of the degree of importance: *it is important/vital/essential*
		Frequency	Opinion of the frequency of the content: *it is common/usual/rare*
		Emotive reaction	Emotive reaction to the content: *it is puzzling/shocking/outrageous*
		Responsibility	Opinion about the responsibility or cause for the content: *it is accident, it is coincidence*
	Others		*it is said/stated/mentioned*
Functional	**Hedged**		Matrix predicates with words or phrases that withhold the writer's full commitment or tone down the writer's assertion lexical words: *possibly, perhaps, seem/appear* modal verbs: *can, could, may, might, would*
	Emphatic		Matrix predicates with words or phrases that emphasize the force of the writer's assertion lexical words: *definitely, obviously, highly, extremely* modal verbs: *must, should* others: *it's my firm belief*

categories (i.e., *Likelihood* and *Certainty*) under the *Semantic* classification. The differences are that the *Epistemic* categories pertain to the overall semantic meaning of the writer's stance toward the content of the extraposed clause whereas *Hedged* and *Emphatic* are concerned with how that stance is presented in the matrix predicate. For example, the following three matrix predicates with a similar semantic meaning were coded as follows:

> *It is appropriate that* ⋯ (*Attitudinal > Evaluative > Appropriateness*)
> *It seems appropriate that* ⋯ (*Attitudinal > Evaluative > Appropriateness, Hedged*)
> *It is highly appropriate that* ⋯. (*Attitudinal > Evaluative > Appropriateness, Emphatic*)

Second, although Table 2 may seem to suggest that certain lexical items are always linked to one and the same categories, a number of predicates were found to be polysemous and classified into different categories depending on the contexts in which they occurred. For example, "it is not *possible* to lessen the emotional turmoil" and "it is *possible* that the worst in the debate on euthanasia is yet to come" were coded as *Attitudinal* (*Dynamic > Potentiality*) and as *Epistemic* (*Likelihood*) respectively.

To improve the reliability of data coding, classification of the instances was conducted together with another applied linguist, a colleague of the researcher, using the coding scheme as presented in Table 2. Inter-coder agreement was 87.1% and 97.7% for the semantic and functional classification respectively. Each discrepancy was resolved after discussion between the two coders.

Finally, log-likelihood[3] was used for significance testing in comparing overall frequencies of the introductory *it* construction and its subtypes in both corpora. As the two corpora are of approximately the same size, the frequency figures provided in the tables in the following Results and Discussion section are not normalized for comparison.

4. Results and Discussion

---

3) Log-likelihood is often used to determine whether the greater frequency of a linguistic item (e.g., a word or multi-word unit) in one corpus than another is statistically significant (Jones & Waller, 2015). In other words, it determines whether the item is statistically overused or underused in one corpus relative to the other.. For specific instructions for computing log-likelihood values, see Jones & Waller (2015).

4.1. Frequencies of the introductory it construction

NICKLE writers turned out to have used the construction significantly more frequently than their NS counterparts. In NICKLE, it occurred 358 times (2.41 per 1,000 words) while there were 265 instances (1.73 per 1000 words) observed in LOCNESS (see Table 3 below). This frequency of use by the Korean writers was also much higher compared to the corresponding figures from corpora of journal articles (0.99) and student dissertations (1.77) in Hewings & Hewings (2002). In terms of syntactic types of the extraposed clauses, *to*-infinitives and *that*-clauses took up a vast majority in both corpora, as was the case with previous studies (Herriman, 2000; Hewings & Hewings, 2002; Zhang, 2015). However, the proportions of the two major syntactic types within each group of writers were considerably different as can be seen in Table 3. Instances of the construction with an extraposed *to*-infinitive subject made up about 65% in NICKLE whereas it was 54% in LOCNESS. In fact, the difference in the overall frequencies of the construction were almost entirely due to the extensive use of *to*-infinitives by the Korean university students.

Table 3. Frequencies of the introductory *it* construction across the two corpora

Syntactic pattern		NICKLE Frequency (% of the total)	LOCNESS Frequency (% of the total)	LL	Sig.	+/−
Non-finite	*to-*	233 (65.1)	142 (53.6)	21.66	***	+
	-ing	2 (0.6)	0	0.00		
Finite	*that-*	118 (33.0)	119 (44.9)	0.00		
	wh-	5 (1.4)	4 (1.5)	0.00		
Total		358	265	13.72	***	+

Note. LL = log likelihood; Sig. = significance (* *p*<0.05, ** *p*<0.01, *** *p*<0.001); "+" indicates overuse in NICKLE and "−" indicates underuse

A couple of possible reasons for this high frequency of the introductory *it* + *to*-infinitive structure may be speculated. First, language proficiency might be at work. Given the choice between *to*-infinitives and *that*-clauses, the Korean writers might have preferred *to*-infinitives as the structure is syntactically easier to use without having to control verb tense or subject-verb agreement. Indeed, this was also observed with student writers in Römer (2009), where the proportion of the

to-infinitive structure was the highest with the less advanced students (73.2%), and the lowest with the NS students (57.2%). Another plausible explanation can be drawn from the possible interaction between the syntactic patterns of the extraposed subjects and stance meanings expressed in the matrix predicates. In other words, the Korean writers may have expressed more often than their NS counterparts certain types of stance that tend to occur with a *to*-infinitive clause, which is discussed in the sections that follow.

4.2. Semantic classification of the matrix predicates

Out of 358 instances of the construction in NICKLE, nearly three quarters (258, 72.1%) were used to express attitudinal stance. As can be seen in Figure 1 below, this proportion is far larger than that of LOCNESS, which was about 57%.
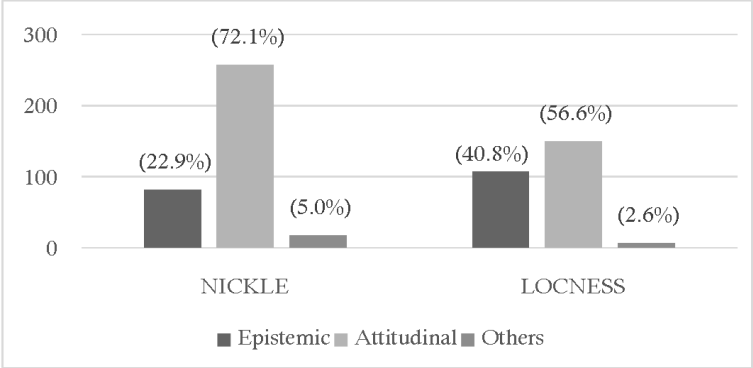

Figure 1. Distribution of semantic types of stance expressed in the introductory *it* construction

By subcategories, the Korean university students used the construction more frequently in all subcategories of *Attitudinal* stance except *Responsibility* but used it less frequently in all *Epistemic* subcategories (see Table 4 below). In what follows, major categories are discussed with actual instances from the two corpora.

Epistemic stance. As shown in Table 4, the Korean university students overall used the introductory *it* construction less to express epistemic stance compared with their NS counterparts although the differences were not great as shown in the relatively small log-likelihood values. In previous studies, epistemic stance was observed in the construction increasingly more toward the formal end of text registers and higher

Table 4. Semantic distribution of stance across the two corpora

	NICKLE	LOCNESS	LL	Sig.	+/−
Epistemic	**82**	**108**	3.63		−
Likelihood	31	50	4.54	*	−
Certainty	51	58	0.47		−
Attitudinal	**258**	**150**	28.68	***	+
<u>Deontic</u>	<u>33</u>	<u>20</u>	3.19		+
Obligation	23	18	0.60		+
Volition	10	2	5.80	*	+
<u>Dynamic</u>	<u>78</u>	<u>47</u>	7.70	**	+
Potentiality	63	36	7.39	**	+
Circumstances	15	11	0.61		+
<u>Evaluative</u>	<u>147</u>	<u>85</u>	16.63	***	+
General evaluation	28	9	10.19	**	+
Appropriateness	64	38	6.64	**	+
Significance	37	22	3.82		+
Frequency	8	7	0.06		+
Emotive reaction	9	3	3.13		+
Responsibility	1	4	1.93		−
Others	**18**	**7**	4.98	*	+

Note. LL = log likelihood; Sig. = significance (* $p<0.05$, ** $p<0.01$, *** $p<0.001$); "+" indicates overuse in NICKLE and "−" indicates underuse

levels of academic writing proficiency (e.g., Gray & Biber, 2012; Römer, 2009; Zhang, 2015). Similarly, in broader EAP research, a high frequency of hedging or stance of likelihood and tentativeness in general has been identified as a feature of mature academic writers (e.g., Hinkel, 2005; Hyland & Milton, 1997). To interpret the result only on this basis, it may be said that NICKLE essays were slightly more informal than their LOCNESS counterparts. Indeed, the Korean university students' underuse of epistemic stance was most pronounced in *Likelihood*. In this category, they not only used a smaller number of tokens, but also relied on only a few types of lexical items such as *seem* and *possible* (for example, the two predicates accounted for about 65% of the total tokens in *Likelihood* as shown in Table 5).

(4) *It seems that* the advertiser think the typical and best role of woman is still a housewife ⋯ (NICKLE, 150)

Table 5. Top three predicates in major semantic categories

	NICKLE	LOCNESS
Epistemic		
*Likelihood	**31**: *seem*(13), *possible*(7), *likely*(3)	**50**: *seem*(10), *possible*(8), *argued*(6)
Certainty	**51**: *true*(13), *certain*(6), *known*(6)	**58**: *obvious*(9), *true*(7), *clear*(4)
Attitudinal		
Dynamic		
*Potentiality	**63**: *hard*(20), *difficult*(14), *easy*(12)	**36**: *difficult*(11), *easy*(9), *possible* (7)
Evaluative		
*General evaluation	**28**: *better*(7), *good*(6), *helpful*(4)	**9**: *better*(2), 8 items occurring once
*Appropriateness	**64**: *natural*(13), *enough*(4), *improper*(4)	**38**: *fair*(5), *okay*(3), *wrong*(3)
Significance	**37**: *important*(32), 5 items occurring once	**22**: *important*(16), *essential*(2), *vital*(2)

Note. * indicates a statistically significant difference between the two groups. Numbers in bold are total numbers of tokens while numbers in parentheses represent the frequencies of the individual predicates.

In contrast, in LOCNESS likelihood stance meanings were expressed with a greater variety of lexical items. Another striking difference found in LOCNESS was that passive verbs such as *argued*, *assumed*, and *suggested* took up a large share (about 40%) of the total instances in the category. In NICKLE, no instances of passive verbs like Example (5) below were observed in the *Epistemic* category. The patterns of passive voice introductory *it* construction are discussed in more detail in section 4.3 below.

(5) However, *it can be argued* that firms in the British economy will suffer from European competition, or even dominance. (LOCNESS, BRSUR3_6)

Attitudinal stance. As noted earlier, NICKLE writers used the introductory *it* construction to express attitudinal stance far more frequently than their LOCNESS counterparts. This was particularly marked in *Potentiality*, *General Evaluation*, and *Appropriateness* (see Table 5). A close look at instances in these categories also reveals

the pattern of a limited range of lexical items taking up a vast majority of the tokens. Indeed, the higher use by NICKLE writers in many categories were simply due to high frequencies of these words. For example, in *Potentiality*, the top three items *hard*, *difficult*, and *easy* made up nearly 75% of the predicates in the category as shown in Table 5.

(6) *it is hard to* find some advertisements with all different races in them. (NICKLE, 41)

This pattern was particularly true of the category *Significance*, which is almost entirely dominated by the word *important*, the most frequent word of all matrix predicates from NICKLE.

(7) *it is important that* presidential candidates let voters know about their policies ⋯ (NICKLE, 83)

These findings are consistent with those from Ädel & Erman (2012), who identified introductory *it* with relatively informal lexical items like *hard* and *easy* as a feature of NNS writers, and Larsson (2017), who observed lower-proficiency students' overreliance on *lexico-grammatical teddy bears*, or a small set of frequent patterns such as *it is important to* and *it is interesting to*.

4.3. Hedged and emphatic stance marking

The Korean university students were found to present their stance in the introductory *it* construction with a hedge to a much less extent than their NS counterparts while using more emphatics as part of the predicates. Table 6 below shows the frequencies and percentages of hedged and emphatic instances of the construction across the two corpora.

Table 6. Hedged and emphatic instances of the introductory it construction

	NICKLE	LOCNESS
Hedged	31 (8.7%)	43 (16.2%)
Emphatic	37 (10.3%)	22 (8.3%)

While their use of hedges and emphatics was not very frequent in both groups, it is suggested that LOCNESS writers tended to present their stance more cautiously while the NICKLE writers tried to accentuate their stance to a greater extent. These results were consistent with Hewings and Hewings (2002), where the student writers seemed to make their arguments more forceful with greater use of emphatics and less use of hedges than the published researchers. It is also in line with the findings from broader EAP studies on NNS argumentative writing that NNS texts tend to have more features of overstatement and involvement than their NS counterparts (Hinkel, 2005; Gilquin & Paquot, 2008). In addition to the higher frequency, some of the emphatics frequently used by the Korean writers were informal lexical items more typically associated with conversational registers such as *very*, *so* and *really* (Biber, 1988) as shown below.

(8) *It was so hard to* write a short paragraph in German and I was expected to get the worst grade. (NICKLE, 109)

These instances show that the Korean writers ironically used highly informal lexical items in a syntactic pattern intended to give an appearance of formality and objectivity. Another noteworthy difference between the two groups in the use of hedges and emphatics was observed in their use of agentless passive voice. While the introductory *it* construction in passive voice (51 instances, 19.2% of the total) occurred more frequently in LOCNESS than in NICKLE (41, 11.5%), many of the passive verbs were communication verbs such as *say*, *report*, and *state*, which, in and of themselves, do not impart any particular stance. However, in NICKLE, many of them were used unmodified as a means of general attribution seemingly in order to add the appearance of generality while hardly expressing any particular stance as illustrated in Example (9) below. Despite the attempt to appear detached, however, the agentless passive sentence may appear vague and inaccurate as there was no further elaboration on or support for the credibility of the statement. In fact, the pattern *it BE said that* were relatively frequent (11 instances), each in a different essay of NICKLE. They were all classified into *Others* in the semantic classification (see Table 4).

(9) *It is said* that there were two big events which changed the view of games greatly ⋯ (NICKLE, 161)

By contrast, most passive verb predicates in LOCNESS were used with either a hedge or emphatic, thereby delivering a stance that can be attributed to the writer. Here, the source of the stance is clear but concealed with the use of agentless passive form as in Examples (10) and (11) below:

(10) *It can be said* that if a single Europe was created, problems may arise over its fundamental political concepts. (LOCNESS, BRSUR3_13)

(11) *It must be noted* that so many Britons regard themselves as being somehow "different" from other Europeans ···(LOCNESS, BRSUR3_6)

These were classified into either *Likelihood* or *Certainty*. This difference shows that LOCNESS writers, if not all, effectively used the construction for its intended rhetorical function of depersonalizing author stance.

4.4. Matching between the semantics of the matrix predicate and the clausal type of extraposed subject

The university writers in both corpora predominantly used non-finite extraposed clauses for expressing attitudinal stance while using mostly finite clauses for marking epistemic stance as seen in Figure 2 below. This is consistent with the findings of the previous studies where there were distinctive associations between attitudinal predicates and extraposed *to*-infinitive clause on one hand, and between epistemic predicates and extraposed *that*-clause on the other (Herriman, 2000; Römer, 2009; Zhang, 2015). This suggests that the Korean writers to some degree appropriately matched stance semantics and the syntactic types of extraposed clauses. The great predominance of attitudinal stance in NICKLE resulted in the frequent use of extraposed *to*-infinitive clauses. However, a close look at the uses of finite extraposed clauses also reveals a considerable difference between the two groups. Whereas finite clauses in both corpora occurred predominantly with epistemic stance, NICKLE writers used *that*-clauses for attitude stance to a much greater extent (30.1%) than their NS counterparts (16.3%). The attitudinal predicates of *that*-clauses in NICKLE were mostly concentrated in *Evaluative* as illustrated in Example (12):

(12) *it is very common that* the stronger try to beautify their history and distort the weaker's history as an uncivilized one (NICKLE, 149)
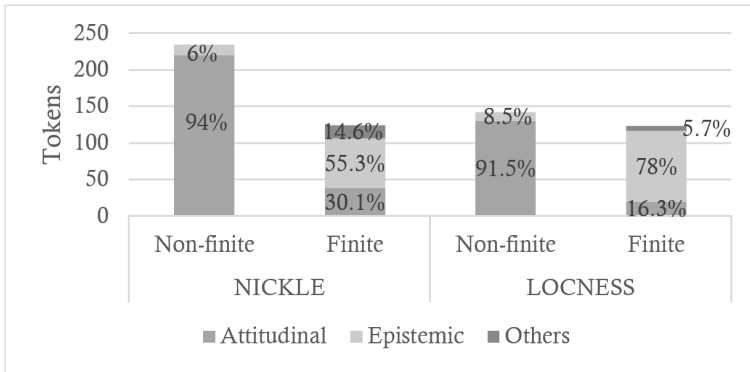
Figure 2. Clausal types of extraposed subjects

This may again indicate the overwhelming dominance of attitudinal stance in NICKLE. However, a detailed perusal also revealed a considerable number of instances where the predicate and the type of extraposed clause was unconventionally or erroneously paired. For instance, *it is common that*-clause in Example (12) above is not ungrammatical but its occurrences are relatively infrequent in authentic texts. It is more common to say '*it is common for* ··· *to*-infinitive'[4]. This mismatch between the two constituents of the construction by the Korean university student writers are discussed in greater detail in the next section along with their other infelicitous uses of the construction.

4.5. Unconventional or erroneous uses of the construction

Results presented above show that the Korean university students represented in NICKLE overall used the introductory *it* construction far more frequently than their NS counterparts to express attitudinal stance with the extraposed *to*-infinitive clause being the most frequent pattern. Among these, however, a number of instances were found to be infelicitous at different lexico-grammatical levels. In what follows, only two major patterns of infelicitous use are presented.

First, the most frequent pattern was to use a predicate that is not normally used with the introductory *it* as illustrated in (13) to (15) below:

---

4) At the time of writing this paper, searches on COCA returned 7 instances of '*it is common that*-clause,' 310 of '*it is common (for* ···*) to*-infinitive'.

(13) *It is sure that* many people older than me also learned English hard. (NICKLE, 243)

(14) *It is no doubt that* if someone wants to be a good man, he or she must be happy first. (NICKLE, 17)

(15) ··· *it is the problem that* there are some advertisements that have implicit false messages like racism or sexism (NICKLE, 5)

This pattern of misuse is widely observed across different writers in NICKLE. These writers may have focused only on the semantics of these lexical items but were not aware of the phraseological restrictions these items have (e.g., *certain* and *sure* are similar in meaning but only *certain* can be used with the introductory *it*). Or they were simply confused with other similarly sequenced phrases, which are either other patterns of the construction such as '*it is without doubt that* ···' and '*it is no doubt true that*' or similar *that*-constructions like '*the problem is that*...'

The second major pattern is a mismatch between the semantics of the predicates and the type of extraposed subject. Example (12) in the previous section belongs to this type of misuse. Here is another example:

(16) *it is very hard that* someone who gets so angry remember the death penalty. (NICKLE, 98)

In Examples (12) and (16), for the extraposed subject of the matrix predicate, the writers chose a *that*-clause rather than an infinitival clause, which would be a far more typical and frequent choice for the respective predicate (*common* and *hard*). Previous studies (e.g., Herriman, 2000) clearly demonstrated that the matrix predicates of *that*-clauses (and *wh*-clauses) mostly represent epistemic stance while those of infinitival clauses represent attitudinal, especially dynamic stance, which was largely confirmed in this study. Herriman (2000) attributed this differing pairing to the semantic differences between finite and non-finite clauses in general. Finite clauses are tensed and typically deliver propositions whose truth value can be assessed (hence, often linked to epistemic stance). By contrast, non-finite clauses mostly represent non-tensed, non-factual and generic actions and states, which can be commented on in terms of potentiality, significance and other attitudinal stance meanings. Given the results discussed in Section 4.4, most NICKLE instances of

the introductory *it* construction overall fit this pattern of pairing between the two elements of the construction. However, there were still quite a few instances of unconventional matching by the Korean university writers.

One notable pattern observed in this type of unconventional use is a missing "linking *to*-infinitive" that would make the content of *that*-clause semantically compatible with the stance meanings in the predicate. The pattern is clearly illustrated in the following examples:

(17) *it is reasonable* that the military life involves some restrictions (NICKLE, 123)

(18) *It does not seem to be logical that* particles are no longer solid objects in the subatomic world (NICKLE, 253)

Searches on the Corpus of Contemporary American English (COCA) for instances of these adjectives (*reasonable* and *logical*) used in the construction show that they are typically followed by a *that*-clause but are linked to it by a *to*-infinitive of mostly communication or cognition verb like *say*, *think*, and *expect* as shown in examples from LOCNESS below:

(19) *It is also reasonable to say that* even though a person afflicted with AIDS may not be the most lucrative investment, we are still talking about a living and breathing human being ⋯ (LOCNESS, USARG_64)

(20) *It would be logical to think that* the leaders of the world would confer and aspire to put a stop to nuclear use ⋯ (LOCNESS, USARG_108)

Without such an infinitival link, the stance expressed in the matrix predicates of Examples (19) and (20) would not be semantically compatible with the content of their extraposed subjects. All these patterns of misuse may indicate that acquisition of the usages of the introductory *it* construction takes multiple levels of lexico-grammatical knowledge, including the one that involves a choice between finite and non-finite clauses.

4.6. Pedagogical implications

The use of the introductory *it* construction by Korean university EFL students as represented in NICKLE proved to be far more frequent than that by their NS counterparts, which is consistent with the results from some previous studies. This may suggest that the construction is quite familiar to Korean EFL university students and conceptually not difficult for them to use. This frequent use of the construction should not be so surprising, however. In fact, a quick look through several textbooks, and English grammar learning websites for Korean secondary school EFL learners indicates that the construction is treated as one of the key syntactic patterns to learn, introduced at relatively early levels. Yet, the treatment of the construction in those books and teaching materials are in large part confined to the end-focus and end-weight principles while rhetorical motivations for using the structure (e.g., depersonalized stance marking) are hardly discussed. Furthermore, the choice between finite and non-finite extraposed clauses are often explained away as a simple stylistic matter suggesting the two types are interchangeable regardless of the semantics of the predicate. Unconventional uses of the introductory *it* construction discussed above may then be traced back in part to the typical ways in which the construction is treated and taught in Korea as well.

Based on these overall results, a number of suggestions can be made to help Korean learners of English, and more broadly EFL academic writers to better understand the construction and use it more appropriately in their writing. At a broad level, the greater frequency of attitudinal stance and emphatics in the construction in NICKLE can be understood in line with the overall tendency of NNS learners' essays to be more informal and involved often with greater use of linguistic features of overstatement (Biber et al., 1999; Hinkel, 2005). While novice EFL academic writers can be guided to increase their awareness of and familiarity with register characteristics of academic writing in general, they can be provided with greater exposure to and explicit instruction in how stance can be expressed in different lexico-grammatical features including the introductory *it* construction (see Biber et al. 1999 for various lexico-grammatical devices for marking stance) and how as a means of negotiating the acceptance of an argument, epistemic stance of expressing doubt and tentativeness and its cautious or guarded presentation are more preferred in academic writing (Gray & Biber, 2012; Hyland & Milton, 1997).

When it comes to the use of the introductory *it* construction, results of this study suggest several specific areas of difficulty in which learners need help. First, learners

should be allowed more opportunities to explore the rhetorical functions of the construction to regulate the writer's commitment to the propositional content and to depersonalize authorial stance in authentic academic texts and to practice using them in their own writing. Particularly, learners can be guided to focus on how to use hedges and agentless passive voice effectively to realize those rhetorical functions. Second, EAP and writing instruction may help learners better understand that a choice between *to*-infinitive and *that*-clause as an extraposed subject is not simply a matter of style by sensitizing them to the semantic differences between them and their associations with specific types of stance. Third, learners should be reminded to avoid repeated or habitual use of a small set of predicates (e.g., *it is important, it is hard,* and *it is said*). They can be encouraged to expand their lexical repertoire for use with different types of stance. Teachers may provide the learners with lexical items that belong to each semantic group of stance but students can also be encouraged to explore for themselves different lexical items and the typical lexico-grammatical contexts in which they are used by accessing resources available online such as phrase banks or corpus tools. Lastly, learners can also be encouraged to extend their linguistic repertoire of syntactic patterns by introducing other patterns that perform rhetorical functions of implicitly marking authorial stance such as stance nouns (Biber et al., 1999, Charles, 2007). All in all, learners should be guided to gradually grow out of lexico-grammatical teddy bears and acquire varied and nuanced ways to mark their stance and express their authorial voice.

5. Conclusions

This study investigated stance marking in the introductory *it* construction by Korean EFL university students as compared to that by NS counterparts, using an integrative classification framework that covers both stance meanings and ways in which they are presented. The study generated three major findings. First, while the Korean EFL writers used the construction far more frequently than NS students in their argumentative essays, most of the use occurred to mark attitudinal stance. Second, the Korean writers were found to present their stance in the construction more emphatically and less cautiously than their NS counterparts. Finally, there were a number of unconventional and erroneous instances of the construction by the Korean students involving the choice of matrix predicate and its semantic compatibility with the syntactic type of the extraposed clause. These findings suggest

that despite their frequent use of the construction, the Korean EFL novice academic writers do not fully utilize the rhetorical functions of the introductory *it* construction. This study thus identified some major areas of difficulty for these learners in the use of the construction and has made some pedagogical suggestions about what can be done in EAP and EFL writing classes to help the learners.

However, the study was not without limitations. Among others, the size of the corpora used was fairly small with both being less than 150,000 words, producing only a small number of instances for some categories. So the characteristics observed of university students' use of the introductory *it* construction in this study cannot be generalizable to broader populations of EFL and non-NS novice academic writers. Another limitation may be that although special care had been taken to make the two corpora as compatible for comparison as possible, there were still differences in terms of essay topics and length, which might have influenced the frequency and distribution of instances in some categories. Finally, as one of the reviewers pointed out, by comparing against the "NS norm," the results of the study may have inadvertently depicted the Korean EFL writers as deficient language users rather than competent multilingual writers with their own intentions and strategies (Canagarajah, 2006).

These limitations suggest possible avenues for future studies. To grasp a more accurate and comprehensive picture of novice academic writers' use of the construction and other lexico-grammatical structures for stance marking, the scope of investigation can be expanded to include writers at different levels of academic writing experience and proficiency using larger corpora with broader and varied perspectives that focus less on the deficiencies of the learners but more on their development and attainment in language use. These investigations can also be conducted with more sophisticated and finer-grained classification frameworks than the one used in the present study.

References

Ädel, A. (2014). Selecting quantitative data for qualitative analysis: A case study connecting a lexicogrammatical pattern to rhetorical moves. *Journal of English for Academic Purposes, 16*, 68-80.

Ädel, A., & Erman, B. (2012). Recurrent word combinations in academic writing by native and non-native speakers of English: A lexical bundles approach. *English for Specific Purposes,*

31(2), 81–92.

Biber, D. (1988). *Variation across speech and writing*. Cambridge, UK: Cambridge University Press.

Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes, 5*(2), 97-116.

Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (1999). *Longman grammar of written and spoken English*. Harlow: Longman.

Canagarajah, S. (2006). Toward a writing pedagogy of shuttling between languages: Learning from multilingual writers. *College English, 68*(6), 589-604.

Charles, M. (2007). Argument or evidence? Disciplinary variation in the use of the noun that pattern in stance construction. *English for Specific Purposes, 26*(2), 203-218.

Collins, P. (1994). Extraposition in English. *Functions of Language, 1*, 17-24.

Gilquin, G. & Paquot, M. (2008). Too chatty: Learner academic writing and register variation. *English Text Construction, 1*(1), 41-61.

Gray, B., & Biber, D. (2012). Current conceptions of stance. In K. Hyland & C. S. Guinda (Eds.), *Stance and voice in written academic genres* (pp. 15-33). New York: Palgrave Macmillan.

Hasselgård, H. (2009). Thematic choice and expressions of stance in English argumentative texts by Norwegian learners. In K. Aijmer (Ed.), *Corpora and language teaching* (pp. 121-139. Amsterdam/Philadelphia: John Benjamins.

Herriman, J. (2000). Extraposition in English: a study of the interaction between the matrix predicate and the type of extraposed clause. *English Studies*, *81*, 582-599.

Hewings, M., & Hewings, A. (2002). "It is interesting to note that…": A comparative study of anticipatory 'it' in student and published writing. *English for Specific Purposes, 21*, 367-383.

Hinkel, E. (2005). Hedging, inflating, and persuading in L2 academic writing. *Applied Language Learning, 15*, 29-53.

Hunston, S. (2004) Counting the uncountable: Problems of identifying evaluation in a text and in a corpus. In A. Partington, J. Morley & L. Haarman (Eds.), *Corpora and discourse* (pp. 157-188). Bern: Peter Lang.

Hunston, S., & Thompson, G. (Eds.). (2000). *Evaluation in text: Authorial stance and the construction of discourse*. Oxford: Oxford University Press.

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. London: Continuum.

Hyland, K. (2016). Writing with attitude: Conveying a stance in academic texts. In E. Hinkel (Ed.), *Teaching English grammar to speakers of other languages* (pp. 246-265). New York: Routledge.

Hyland, K., & Milton, J. (1997). Qualification and certainty in L1 and L2 students' writing. *Journal of Second Language Writing, 6*(2), 183-205.

Hyland, K., & Tse, P. (2005). Evaluative that constructions: Signalling stance in research abstracts. *Functions of Language, 12* (1), 39-63.

Jacobs, R. A. (1995). *English syntax: a grammar for English language professionals*. New York: Oxford University Press.

Jiang, F. (2015). Nominal stance construction in L1 and L2 students' writing. *Journal of English for Academic Purposes*, *20*, 90-102.

Jiang, F., & Hyland, K. (2018). Nouns and academic interactions: A neglected feature of metadiscourse. *Applied Linguistics*, 39(4), 508-531.

Jones, C., & Waller, D. (2015). *Corpus linguistics for grammar: A guide for research*. London: Routledge.

Kaltenböck, G. (2000). It-extraposition and non-extraposition in English discourse. In C. Mair & M. Hundt (Eds.), *Corpus linguistics and linguistic theory* (pp. 157-175). Amsterdam: Rodopi.

Larsson, T. (2017). A functional classification of the introductory it pattern: Investigating academic writing by non-native-speaker and native-speaker students. *Journal of English for Academic Purposes, 48*, 57-70.

Martin, J. R., & White, P. R. R. (2005). *The language of evaluation: Appraisal in English*. Basingstoke, UK: Palgrave Macmillan.

Nini, A. (2014). Multidimensional analysis tagger (Version 1.2)[Computer software]. Retrieved from http://sites.google.com/site/multidimensionaltagger

Peacock, M. (2011). A comparative study of introductory it in research articles across eight disciplines. *International Journal of Corpus Linguistics, 16*, 72-100.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. London: Longman.

Römer, U. (2009). The inseparability of lexis and grammar: Corpus linguistic perspectives. *Annual Review of Cognitive Linguistics, 7*, 140-162.

Wingate, U. (2012). 'Argument!' helping students understand what essay writing is about. *Journal of English for Academic Purposes, 11*(2), 145-154.

Zhang, G. (2015). It is suggested that…or it is better to.? Forms and meanings of subject it-extraposition in academic and popular writing. *Journal of English for Academic Purposes, 20*, 1-13.

Choongil Yoon
Assistant Professor
Paramita College
Dongguk University (Gyeongju Campus)
123 Dongdae-ro, Gyeongju, 38066 Gyeonsangbuk-do, Korea
E-mail: chongal2@hotmail.com