Korean Journal of English Language and Linguistics, Vol 24, October 2024, pp. 1071-1100
DOI: 10.15738/kjell.24..202410.1071

KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

ISSN: 1598-1398 / e-ISSN 2586-7474

http://journal.kasell.or.kr

Sentence-Initial Particularizing Adverbials in EFL Learner Writing: The Role of L1 Background and L2 Proficiency

Seung-Ah Lee (Ewha Womans University)

Received: August 27, 2024
Revised: September 23, 2024
Accepted: September 28, 2024

Lee, Seung-Ah
Professor, Division of English Language and Literature
Ewha Womans University
52, Ewhayeodae-gil, Seodaemun-gu
Seoul, Korea
Tel: 02) 3277-4507
Email: seungahlee@ewha.ac.kr

ABSTRACT

Lee, Seung-Ah. 2024. Sentence-initial particularizing adverbials in EFL learner writing: The role of L1 background and L2 proficiency. *Korean Journal of English Language and Linguistics* 24, 1071-1100.

This paper presents a learner corpus-based study of English particularizing adverbials, a comparatively under-researched topic with regard to adverbial placement. Particularizers such as *especially*, *particularly*, and *in particular* highlight the fact that an utterance is confined to the part that is focused. Pedagogically-oriented sources and corpus-informed dictionaries state that *especially* cannot appear at the beginning of a sentence before the subject. The primary goal of this study is to determine whether misused sentence-initial *Especially* is a production tendency associated with learners' first language (L1) or a universal feature of interlanguage development. To this end, I examined 3,800 argumentative essays in the International Corpus Network of Asian Learners of English (ICNALE), which were written by English as a foreign language (EFL) learners from Korea, China, Indonesia, Japan, Taiwan, and Thailand. The learner participants in the ICNALE are classified according to four English proficiency levels. In order to identify the patterns of overuse and underuse, EFL learners' essays were also compared to 200 essays written by L1 English novice writers in the ICNALE. Both L1 background and English proficiency significantly influenced the frequency of misused sentence-initial *Especially* although the effect size was small in the case of the latter. The relatively high rate of misuse in L1 Korean writers' essays can be attributed to L1 transfer, the effect of which is confirmed by an investigation of a Korean language corpus. The findings suggest that formal L2 instruction based on data-driven learning may be an effective way to remedy the misuse of sentence-initial *Especially*.

KEYWORDS

particularizer, restrictive adverb, adverb placement, second language writing, learner corpus, ICNALE

1. Introduction

A number of corpus-based studies have found that there are differences of adverbial placement between first language (L1) writing and second language (L2) writing. Previous studies have commonly focused on the overuse of initial positioning in L2 English writing. L2 English learners' preference for sentence-initial positions has been predominantly discussed in relation to linking adverbials such as *however*, *therefore*, etc. (Dupont and Granger 2022, Green et al. 2000, Tribble 1989, Van Vuuren and Berns 2018, Zhang 2000). The work by Tribble (1989) was perhaps one of the earliest studies comparing the proportional use of *however* in L1 and L2 corpora. Only 25% of the L1 writers' use of *however* was sentence-initial, while 81% of the 48 instances in the L2 corpus were. Other existing research, such as Osborne (2008), has addressed the placement of frequency adverbs (*often, sometimes*, etc.) in the context of L1-L2 differences of syntactic patterns (i.e., adverb-verb-object order in L2 English vs. verb-adverb-object order in L1s such as French, Italian, and Spanish).

The present study concerns the position of English particularizing adverbials, a hitherto neglected area of learner corpus research. Particularizers such as *especially* highlight the fact that an utterance is confined to the part that is focused. Swan (2005) and several corpus-informed dictionaries draw attention to the fact that *especially* cannot appear at the beginning of a sentence before the subject. Yet, misused sentence-initial *Especially* (marked by a capital) is found in Korean English as a foreign language (EFL) student writing, which motivated this study. All examples in (1) were produced by different undergraduate students taking an English grammar class at a university in Seoul, who were asked to write a peer review for other students' class presentations.

(1) a. Especially, I liked the introduction with daily examples and a witty cartoon.
b. Especially, I liked the contexts that are from the presenter's real life experience.
c. Especially, the second part of the presentation that dealt with […] was very impressive.
d. Especially, she gave us exercise question, so I could have some time to think about the answer.
e. Especially, she presented not only the subjunctive mood but also the indicative mood, which helped me to understand […].

Such non-target-like placement of sentence-initial *Especially* is not infrequent even in the writings of highly advanced Korean EFL learners. Specific examples, which are mostly from published journal articles on English linguistics or English teaching, can be provided upon request.

The non-target-like placement of sentence-initial *Especially* in (1) clearly differs from other cases of initial positioning. For one thing, initially-placed *However* is not ungrammatical or unacceptable. For another, sentence-initial *And* and *But*, which have been proscribed by prescriptive grammarians, occur not only in L2 writing but also in L1 writing, though sentence-initial *And* is much less frequent than sentence-initial *But* (Bell 2007, Pyo 2018).

A question that needs to be investigated is whether misused sentence-initial *Especially* can be attributed to intralingual or developmental errors (Richards 1974). That is, is this interlanguage phenomenon caused by learners' ignorance of the target language feature (namely, the fact that unlike other particularizers in English, *especially* cannot modify an entire clause)? A different analysis is that the interlanguage phenomenon in (1) can be associated with L1 transfer. Although the primary focus of this study is on Korean EFL learner writing, simply relying on mono-L1 learner corpora, such as the Yonsei English Learner Corpus (YELC, Rhee and Jung 2014) or the Gachon Learner Corpus (Carlstrom and Price 2012-2014), is not sufficiently helpful to argue for one position

over the other. Thus, by employing a multi-L1 learner corpus, the present study aims to answer the following research questions:

1. What is the frequency distribution of *especially* and other synonymous particularizers used in sentence-initial and non-sentence-initial positions of L1 and L2 writing? Can overuse/underuse patterns be discovered?
2. How is sentence-initial *Especially* used by L1 English writers and EFL writers, respectively?
3. Do learners' L1 background and L2 proficiency affect the frequency of misused sentence-initial *Especially* in EFL learner writing?

2. English Particularizers

Whereas the term 'adverb' refers to a single word, the term 'adverbial' includes a group of words such as a prepositional phrase. Biber et al. (1999) classify adverbials into three main categories according to their functions. The first of these is called circumstance adverbials (e.g., *just*, *only*, *especially*, *particularly*, *in particular*, and *also*), which have the primary function of adding information about the proposition expressed in the clause. The second category is called stance adverbials (e.g., *really*, *actually*, *of course*, *perhaps*, and *probably*), which convey the stance of the speaker/writer toward the clause. The third category is called linking adverbials (e.g., *so*, *thus*, *therefore*, *however*, and *for example*), which mark the relationship between two pieces of discourse.

Circumstance adverbials fall into a variety of subcategories and further subcategories. According to Biber et al. (1999), restrictive circumstance adverbials such as *only*, *especially*, and *in particular* "emphasize that the proposition is true in a way which expressly excludes some other possibilities", while additive circumstance adverbials such as *also* and *too* "show that a current proposition is being added to previous one" (p. 780). Although there is a slight discrepancy in the taxonomy and nomenclature of circumstance adverbials, the division between restrictive and additive adverbials is widely accepted by grammarians including Quirk et al. (1985) and Huddleston and Pullum (2002), as shown in Tables 1 and 2.

A characteristic shared by both restrictive and additive adverbials is that their position is crucial in deciding what element of the clause is the focus of the restriction or addition. That is, unlike many other adverbials, restrictive and additive adverbials "often cannot be moved without affecting their meaning in the clause" (Biber et al. 1999, p. 781), as illustrated in (2)-(3). (2a) is taken from a fiction text (FICT) and (3a) is taken from a newspaper text (NEWS). [1]

(2) Restrictive
  a. A heart born ***especially*** for me, Jackie used to tease. (FICT)
  b. ***Especially*** a heart born for me, Jackie used to tease.

(3) Additive
  a. Mr. Arce Gomez ***also*** has a grisly human rights reputation. (NEWS†)
  b. ***Also*** Mr. Arce Gomez has a grisly human rights reputation.

(Biber et al. 1999, p. 781)

[1] The dagger icon (†) indicates that (3a) is truncated for space reasons.

Table 1. Quirk et al.'s (1985, p. 604) Two Main Subdivisions of Focusing Subjuncts

Category	Subcategory	Description	Examples
Restrictive Subjuncts		The utterance concerned is true in respect of the part focused.	
	Exclusives	They restrict the application of the utterance *exclusively* to the part focused.	*alone, exactly, exclusively, just, merely, only, precisely, purely, simply, solely*
	Particularizers	They restrict the application of the utterance *predominantly* to the part focused.	*chiefly, especially, largely, mainly, mostly, notably, particularly, primarily, principally, specifically; at least, in particular*
Additive Subjuncts		The utterance concerned is *additionally* true in respect of the part focused.	*again, also, either, equally, even, further, likewise, neither, nor, similarly, too; as well, in addition*

Table 2. Huddleston and Pullum's (2002, p. 587, p. 592) Two Main Subdivisions of Focusing Modifiers

Category	Subcategory	Description	Examples
Restrictive Focusing Modifiers	Total Restrictive Focusing Modifiers	The restriction expressed by them is total.	*alone, but, exactly, exclusively, just, merely, only, precisely, purely, simply, solely*
	Partial Restrictive Focusing Modifiers	The restriction expressed by them is partial.	*chiefly, especially, mainly, mostly, notably, particularly, primarily; at least, for the most part, in particular*
Additive Focusing Modifiers			*also, even, too; as well, in addition*

Examples in (2) illustrate fronting (Biber et al. 1999) or complement preposing (Huddleston and Pullum 2002), in which the noun phrase is placed in the initial position of a sentence. In (2a), the focus is *me* or (equivalently) *for me*, whereas in (2b), it is the noun phrase *a heart born for me*. In (3a), *also* can have a variety of foci, such as *Mr. Arce Gomez, a grisly human rights reputation*, or *has a grisly human rights reputation* (cf. Huddleston and Pullum 2002, p. 593). Additionally, *also* can occur in the sentence-initial position, with the whole clause as the focus, as in (3b).

Following Quirk et al. (1985) and Nelson et al. (2002), in this paper, I will use the term 'particularizer' to refer to expressions such as *especially*, *particularly*, *in particular*, and so on. In the British Component of the International Corpus of English (ICE-GB), particularizers are tagged 'ADV(partic)' (Nelson et al. 2002). The adverb *specially* is tagged 'ADV(partic)' in the ICE-GB, although it is not included in Tables 1 and 2.

Among the various English particularizers, *especially* deserves special attention with regard to its position. Swan (2005) and several corpus-based monolingual dictionaries provide the following minimal pairs of grammatical and ungrammatical sentences:

(4) a. All my family like music. My father, especially, goes to as many concerts as he can.
b. *Especially my father goes to as many concerts as he can.

(Swan 2005, p. 169)

(5) a. I am especially grateful to all my family and friends who supported me.
b. *Especially I am grateful to all my family and friends who supported me.
(https://dictionary.cambridge.org/grammar/british-grammar/especially-or-specially)

(6) a. I especially like sweet things.
b. *Especially I like sweet things.
(https://www.oxfordlearnersdictionaries.com/definition/english/especially?q=especially)

(7) a. I like Thai food especially.
b. I especially like Thai food.
c. *Especially I like Thai food.
(https://www.ldoceonline.com/dictionary/especially)

(8) a. Young people especially are being affected by the economic crisis.
b. *Especially young people are being affected by the economic crisis.
(https://www.ldoceonline.com/dictionary/especially)

In all the ungrammatical sentences in (4)-(8), sentence-initial *Especially* precedes the subject of a clause. This restriction is expressed in different wordings, as listed in (9).

(9) a. *Especially* follows a subject.
(Swan 2005, p. 169)
b. No explanation was given.
(https://dictionary.cambridge.org/grammar/british-grammar/especially-or-specially)
c. *Especially* is not placed first in a sentence.
(https://www.oxfordlearnersdictionaries.com/definition/english/especially?q=especially)
d. *Especially* is not used at the start of a sentence before the subject.
(https://www.ldoceonline.com/dictionary/especially)

In order to demonstrate that the explanation in (9c) is not sufficient enough, examples drawn from a large native corpus, namely the Corpus of Contemporary American English (COCA, Davies 2008-), are given in Table 3. In Table 3, the category 'other conjunction' refers to conjunctions excluding *if* and *when*, the two most frequent conjunctions after *especially*. The category 'complex preposition' includes expressions such as *thanks to*, *due to*, *because of*, etc. The COCA is a balanced corpus and is evenly divided between the following eight genres: blogs (BLOG), web pages (WEB), TV and Movies subtitles (TV/M), spoken (SPOK), fiction (FIC), popular magazines (MAG), newspapers (NEWS), and academic journals (ACAD). As shown in (9), when considering the position of *especially*, it is crucial to pay attention to the subject of a clause. Hence, in the case of the category 'subject', three examples, each from a different genre, are provided in Table 3.

According to Biber et al. (1999), *especially* and *particularly* occur "at least 200 times per million words" in academic prose (pp. 561-562). Figure 1 shows the distribution of *especially* (case-insensitive) in the eight genres of the COCA (retrieved on August 19, 2024).[2] The information obtained by using the COCA Chart function is

[2] The new version of the COCA is released in March 2020. The COCA website (https://www.english-corpora.org/coca/) seems

read as follows. First, 'FREQ' shows the raw frequency of *especially* in each register (i.e., section). Second, 'WORDS (M)' refers to the size of each section in millions of words. For example, the COCA academic section contains 119,800,000 words. Third, 'PER MIL' indicates the frequency of *especially* per million words.

Table 3. Positional Variation of *Especially* in the COCA

Position	Neighboring Constituent	Example	Genre
Initial	+ PP	***Especially** in a small room, it's fabulous.*	BLOG
	+ Adverb	***Especially** now, every point matters.*	NEWS
	+ *If/When*-Clause	***Especially** if you have never painted before, there are my top five things recommended to get you off on the right foot.*	WEB
	+ Other Conjunction	***Especially** because I am a senior. I can help my team.*	BLOG
	+ Complex Preposition	***Especially** thanks to the rise of BYOD with mobile, we as consumers feel empowered that our choices will shape the industry.*	WEB
	+ Adjective (Inversion)	***Especially** important is the area of informed consent.*	ACAD
	+ Sentence Fragment (NP)	***Especially** children with special needs.*	SPOK
	+ Sentence Fragment (PP)	***Especially** for this reason.*	SPOK
Medial	Subject +	*We **especially** liked that poem.*	NEWS
	Subject +	*Women **especially** were concerned about the effects of militant Islam on advancing democratization.*	ACAD
	Subject +	*This virus **especially** was lethal for young adults, and particularly pregnant women.*	MAG
	+ PP	*The goal is to save money while providing better care, **especially** for people with chronic illnesses.*	NEWS
	+ Adverb	*We've both been so busy, **especially** lately.*	FIC
	+ *If/When*-Clause	*Guys never want to hear that, **especially** when they're already married and have kids.*	TV/M
	+ Other Conjunction	*It increases the complexity of the operations, **especially** since there is no communication across different systems.*	ACAD
	+ Complex Preposition	*But I remember them **especially** because of that grand house.*	FIC
	+ Adjective	*Cassie had always been **especially** close to her father.*	FIC
	+ NP	*This is **especially** the case with ecstasy.*	ACAD
	NP, + NP	*The family, **especially** the children, turned to him from the start with an assumption of solidarity and joy.*	ACAD
	to + VP	*But I want to **especially** thank President Clinton.*	SPOK
Final	PP +	*Some surely results from understandable apprehension about U.S. courses of action, in the Middle East **especially**.*	ACAD
	NP +	*Many cancer-prevention diets recommend broccoli **especially**.*	MAG
	NP, NP +	*Don Sandy has always loved cars, European sports cars **especially**.*	MAG

Note. The '+' symbol indicates the position of *especially*. PP = prepositional phrase, NP = noun phrase, VP = verb phrase.

As shown in Figure 1, *especially* occurs most frequently in the academic genre. Table 4 presents the distribution of six particularizers in the COCA academic section according to their positions: initial, medial, and final (retrieved on August 19, 2024). All six particularizers in Table 4 have a preference for sentence-medial positions. The

to be updated. When accessed the COCA website on July 26, 2024, the numbers in Figure 1 and Table 4 were very slightly different. Thus, the retrieval date is important.

pervasive use of medial positions is mentioned in some grammar books without reference to register distinctions. According to Quirk et al. (1985), *in particular* prefers a position "after the focused part", as in *The workers, in particular, are dissatisfied with the government* (p. 604, p. 608). Biber et al. (1999) also note that restrictive adverbials favor medial positions because these positions "show clearly the scope of the adverbial" (p. 805).

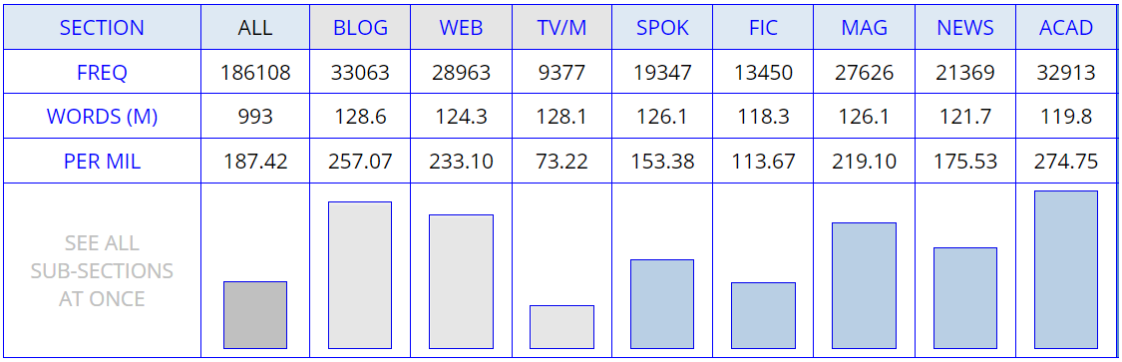

Figure 1. Frequency of *Especially* in the Different Genres of the COCA

Table 4. Frequency (Per Million Words) of Six Particularizers in the COCA Academic Section Across Positions

Position	*Especially*	*Particularly*	*In Particular*	*Specifically*	*Notably*	*Specially*
Initial	4.69	4.27	23.08	31.02	4.61	0.18
Medial	269.88	217.21	56.32	97.39	24.13	6.24
Final	0.18	0.18	5.46	1.26	0.03	0.03
Total	274.75	221.66	84.86	129.67	28.77	6.45

Except for *especially*, corpus-informed dictionaries do not pay attention to the positions of the particularizers shown in Table 4. Thus, I extracted the following sentence-initial examples of these six particularizers from the Cambridge Dictionary (https://dictionary.cambridge.org/):

(10) a. Especially in less developed countries, research on ageing still tends to prejudge people's and to focus narrowly on material outcomes.
b. Particularly this latter section is informed by "behind-the-scenes stories," based on copious interviews with key players in any given project.
c. In particular, we identified the fact that the paths are the only tractable component of the design space.
d. Specifically, we found high rates of psychiatric distress, in the form of depression.
e. Notably, these differences were eventually equalized by the turning-point experience of study abroad.
f. Specially designed materials were used in the experiment.
(From the Cambridge English Corpus, cited in https://dictionary.cambridge.org/)

In (10a), sentence-initial *Especially* is followed by a prepositional phrase. As shown in (10b)-(10e), sentence-initial *Particularly*, *In particular*, *Specifically*, and *Notably* can occur before the subject of a clause, suggesting that they may have scope over the entire clause. Sentence-initial *Specially* is usually followed by a past participle, as in (10f).

As noted in the previous section, the first research question of this study concerns *especially* and other synonymous particularizers. Of the six particularizers mentioned in Table 4, only the first three will be discussed in the remainder of this paper. *Specifically* will not be of further concern in this study because its Korean translation equivalent is *kwucheycekulo* rather than *thukhi* ('especially').[3] *Notably* and *specially* are relatively lower frequency words, and thus are not of primary interest.

3. Methodology

3.1 The Corpus and Subcorpora Used for the Analysis

The corpus used in this study was the International Corpus Network of Asian Learners of English (ICNALE, Ishikawa 2023). As of May 2024, the ICNALE consists of five major modules: Spoken Monologues (SM), Spoken Dialogues (SD), Written Essays (WE), Edited Essays (EE), and Written Essays 2 (WE2). Of these, the present study used Written Essays (WE) version 2.6 (updated in January 2024). Both WE and WE2 modules comprise 200-300-word essays about two ICNALE common topics: (i) a part-time job for college students (PTJ) and (ii) non-smoking at restaurants (SMK). WE2 is the new module collected from new Asian countries (i.e., Bangladesh, Cambodia, India, Laos, Malaysia, and Myanmar). WE2 was not used in this study because the ICNALE learner background survey sheet, available on the ICNALE website (https://language.sakura.ne.jp/icnale/), contains information on SM, SD, WE, and EE only.

All participants in the ICNALE WE were asked to write argumentative essays on the following two topics:

(11) Do you agree or disagree with the following statements? Use reasons and specific details to support your answer.
A: It is important for college students to have a part-time job.
B: Smoking should be completely banned at all restaurants in the country.

(Ishikawa 2011, p. 5)

The conditions of writing are described in Table 5.

Table 5. Writing Conditions of the ICNALE (Ishikawa 2023, p. 23)

Condition	Details
Time for preparation	Included in the time for writing
Time for a task	20 to 40 minutes for one essay
Length	200 to 300 words
Reference use	No
Data collection method	Written on MS Word
Exam condition	No

One of the key features of the ICNALE is that every learner has been assigned to four proficiency bands of the Common European Framework of Reference for Languages (CEFR): A2, B1 Lower (B1_1), B1 Upper (B1_2), and B2+, according to their scores in the standardized English proficiency tests or in the Vocabulary Size Test (VST, Nation and Beglar 2007). Table 6 provides a more detailed explanation about these four proficiency bands.

[3] In this paper, the Yale romanization of Korean is used.

Table 6. Score Conversion Table (Ishikawa 2023, p. 27)

Level	TOEIC L/R	TOEFL PBT	TOEFL iBT	IELTS	VST
A2 (Waystage)	–545	–486	–56	3+	–24
B1_1 (Threshold Lower)	550+	487+	57+	4+	25+
B1_2 (Threshold Upper)	670+	527+	72+	4+	36+
B2+ (Vantage or Higher)	785+	567+	87+	5 (5.5)+	47+

Note. L/R = Listening and Reading test, PBT = paper-based test, iBT = internet-based test.

The proficiency-based classification of non-native writers was the primary reason for selecting the ICNALE rather than other written learner corpora. The third version of the International Corpus of Learner English (ICLEv3, Granger et al. 2020) contains English argumentative essays produced by learners from the following L1s: Brazilian Portuguese, Greek, Hungarian, Persian (Iran), Korean, Lithuanian, Macedonian, Pakistani, and Serbian. ICLEv3, however, does not offer learner metadata relating to English proficiency. The second version of the ICLE (ICLEv2, Granger et al. 2009) covers English learners from the following 16 L1 backgrounds: Bulgarian, Chinese, Czech, Dutch, Finnish, French, German, Italian, Japanese, Norwegian, Polish, Russian, Spanish, Swedish, Turkish, and Tswana. A professional rater classified only a random sample of 20 essays from each of these 16 subcorpora into three CEFR levels: B2 (and lower), C1, and C2 (Granger et al. 2009). In short, although both the ICNALE and the ICLE are currently available, representative multi-L1 learner corpora (Granger 2012, Granger 2013), not all learners in the ICLE were classified into CEFR levels. All we know about their proficiency is that they are higher intermediate to advanced EFL learners.

The ICNALE includes the data of Asian learners not only in EFL regions (China, Indonesia, Japan, Korea, Taiwan, and Thailand) but also in English as a second language (ESL) regions. As the central concern of this study is EFL Korean student writing, ESL regions (Hong Kong, Pakistan, the Philippines, and Singapore) were excluded from the analysis. Throughout this paper, the following abbreviations will be used where necessary for the six EFL regions: Korea (KOR), China (CHN), Indonesia (IDN), Japan (JPN), Thailand (THA), and Taiwan (TWN).

Another advantage of the ICNALE is the inclusion of data produced by L1 English native speakers (ENS). Thus, unlike the case of ICLE, a separate L1 reference corpus is not necessary when comparing L1 and L2 writing. Moreover, the ENS data in the ICNALE can be grouped according to occupational variety: ENS students (college students), ENS teachers (English teachers, instructors, and professors), and ENS others (adults with varied job backgrounds). The L2 learners in the ICNALE are college students. Thus, the subcorpus of ENS students is selected as a "peer reference" and the subcorpus of ENS teachers is regarded as a "pedagogical reference" (Ishikawa 2023, p. 19). The ENS data in the ICNALE can also be categorized into nationalities: the USA, Great Britain (GBR), Australia, Canada, and New Zealand. In Korea, American English is generally considered as a primary learning model. The same is true for Japan, while in (many parts of) mainland China, "a British English model is adopted" (Ishikawa 2023, p. 18). In the hope of finding some useful information, I also created ENS USA and ENS GBR subcorpora.

Table 7 and Figure 2 present an overview of the ICNALE Written Essays subcorpora used in the present study. The number of words, as well as the number of essays (i.e., files), of each subcorpus was calculated using AntConc (version 4.3.0; Anthony 2024). As shown in Table 7, all of the ICNALE participants wrote two essays. In Table 7, the last column is added for convenience by the author. The numbers in this column are intended to facilitate a more direct comparison across subcorpora of different sizes. For example, in the subcorpus of ENS students, one occurrence per million words equals 22, while it is 30 in the KOR A2 subcorpus.

Table 7. Description of the ICNALE Written Essays Subcorpora Used in the Present Study

Subcorpus	Number of words	Number of essays	Number of participants	One occurrence per million words
ENS students	45,028	200	100	22
ENS teachers	20,003	88	44	50
ENS students USA	36,177	164	82	28
ENS students GBR	1,430	6	3	699
ENS teachers USA	5,563	24	12	180
ENS teachers GBR	5,577	24	12	179
KOR A2	33,165	150	75	30
KOR B1_1	26,910	122	61	37
KOR B1_2	40,073	176	88	25
KOR B2+	36,194	152	76	28
KOR	**136,342**	**600**	**300**	**7**
CHN A2	22,520	100	50	44
CHN B1_1	110,289	464	232	9
CHN B1_2	52,031	210	105	19
CHN B2+	6,561	26	13	152
CHN	**191,401**	**800**	**400**	**5**
IDN A2	14,754	64	32	68
IDN B1_1	37,465	164	82	27
IDN B1_2	39,515	166	83	25
IDN B2+	1,541	6	3	649
IDN	**93,275**	**400**	**200**	**11**
JPN A2	68,528	308	154	15
JPN B1_1	79,591	358	179	13
JPN B1_2	22,390	98	49	45
JPN B2+	8,532	36	18	117
JPN	**179,041**	**800**	**400**	**6**
THA A2	53,496	238	119	19
THA B1_1	80,758	358	179	12
THA B1_2	45,829	200	100	22
THA B2+	1,016	4	2	984
THA	**181,099**	**800**	**400**	**6**
TWN A2	12,878	58	29	78
TWN B1_1	40,008	174	87	25
TWN B1_2	28,484	122	61	35
TWN B2+	11,014	46	23	91
TWN	**92,384**	**400**	**200**	**11**
EFL total	873,542	3,800	1900	1
EFL total + ENS students	918,570	4,000	2,000	1

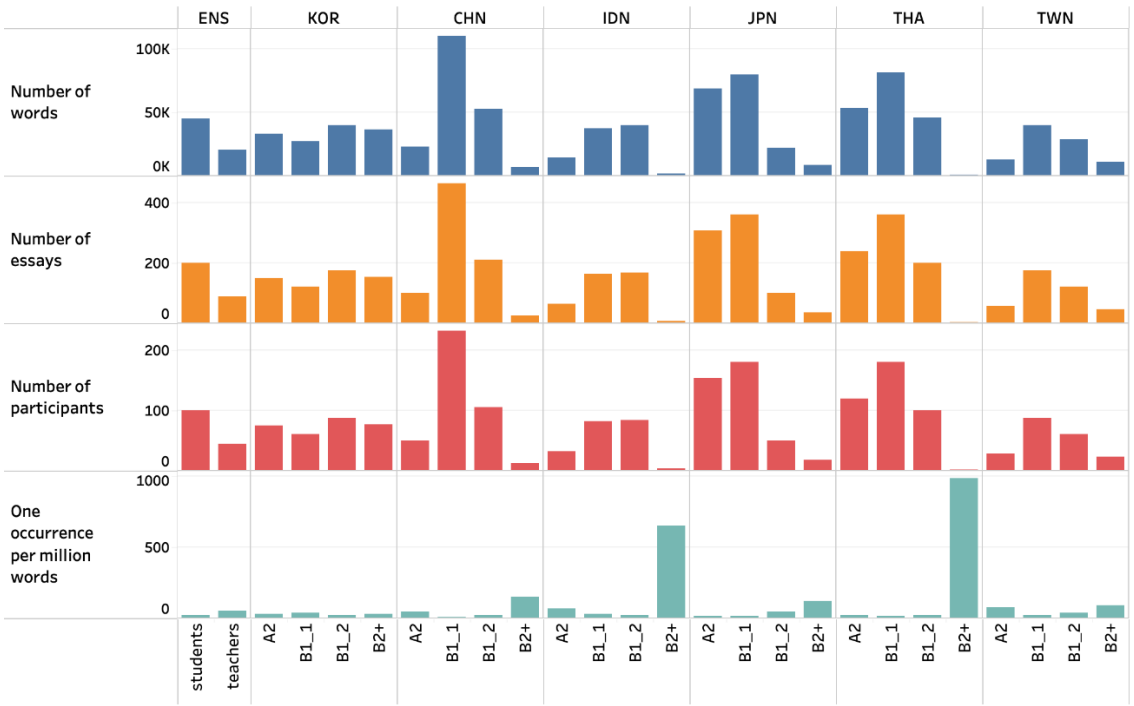

Figure 2. Graphic Overview of the ICNALE Written Essays Subcorpora Used in the Present Study

3.2 Data Collection

This study used the concordancing program AntConc version 4.3.0 to search for instances of particularizers in the ICNALE. Unlike the COCA interface, both the ICNALE Online corpus retrieval system and the AntConc software allow case-sensitive search. After downloading the text (.txt) files of WE version 2.6 from the ICNALE website (https://language.sakura.ne.jp/icnale/), I created individual subcorpora by using the Corpus Manager tool in AntConc. The KWIC (Key-Word-In-Context) tool was used for the corpus searches of particularizers. The search results were saved into Excel files and the obtained corpus data were then coded and annotated by the author.

3.3 Data Analysis

Since the size (i.e., the total number of words) of each subcorpus was different, as shown in Table 7 and Figure 2, I converted raw frequencies (RFs) to normalized frequencies (NFs) for comparison. The conversion was calculated based on the following equation:

(12) NF = RF × the number of one occurrence per million words in each subcorpus (shown in the last column of Table 7)

As represented in (12), the NF in this study refers to the frequency of occurrence per million words, which allows for a direct comparison with the COCA data (see Figure 1 and Table 4). An alternative way of calculating NF is to divide the RF by the total number of words in each subcorpus and multiply the result by one million (Lee 2016). Yet using the formula provided in (12) is more convenient because in this study, the data were coded in spreadsheets such as those shown in Table 10.

In addition to using the download version of the ICNALE, in this research I also used the ICNALE Online (http://language.sakura.ne.jp/onlinecorpus.html) for the purpose of using the Keyword search. The keyword analysis identifies words that appear unusually frequently or unusually infrequently in the target corpus (e.g., KOR A2) in comparison with the words in the reference corpus (e.g., ENS students) based on a statistical measure such as the chi-squared score or the log-likelihood. The Keyword tool is also available in AntConc version 4.3.0. However, the Keyword search in the ICNALE Online is more useful because case sensitivity of particularizers is crucial in the current study. Figure 3 displays the screenshot of an example of the Keyword search in the ICNALE Online. As shown in the screenshot of the results (Figure 4), the default setting is the chi-squared value ("Chi2").

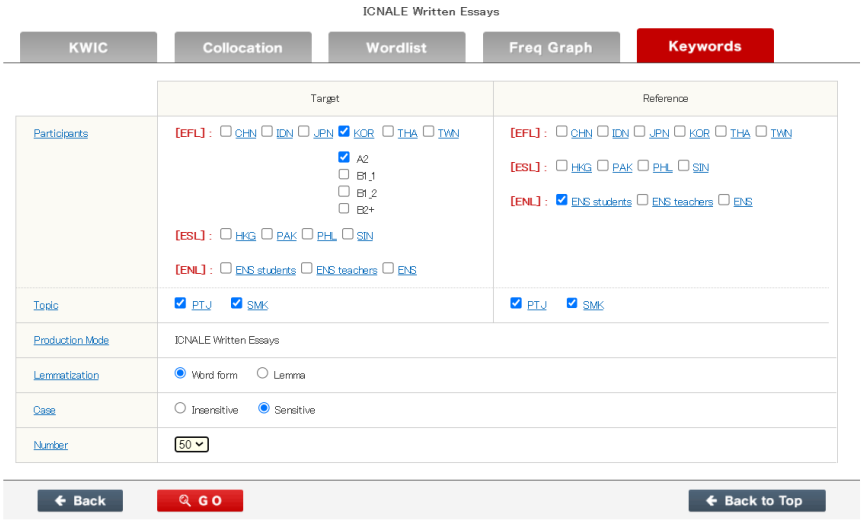

Figure 3. An Example of the Keyword Search in the ICNALE Online

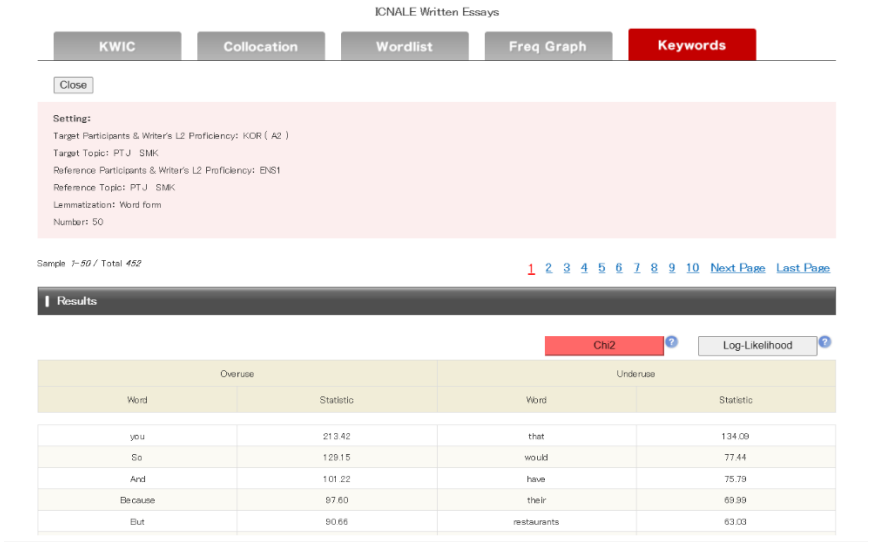

Figure 4. Results of the Keyword Search in Figure 3

Admittedly, there are some minor drawbacks of using the ICNALE Online. First, as shown in Figure 3, only a tripartite classification of the ENS subcorpus is possible: ENS students, ENS teachers, and ENS others. Further subdivisions into nationalities are impossible. Second, the data are currently not updated on the ICNALE Online. The most recently updated WE version 2.6 is available only in the download version. When using the KWIC tool in the ICNALE Online, the raw frequencies of particularizers were exactly the same as those presented in Table 11 in section 4.1. To the best of my knowledge, however, the exact size of each subcorpus is not identifiable in the ICNALE Online. Given the very small differences found in the number of words in WE versions 2.5 and 2.6, presented in Table 8, it is difficult to anticipate any serious problem in using the ICNALE Online, together with the download version.

Table 8. Number of Words in the Different Versions of the ICNALE Written Essays (WE)

Subcorpus	WE Version 2.5 (Updated in June 2023)	WE Version 2.6 (Updated in January 2024)
KOR A2	33,167	33,165
CHN B1_1	110,293	110,289
IDN B1_2	39,517	39,515
JPN A2	68,529	68,528
THA A2	53,506	53,496
THA B1_2	45,840	45,829

In the analysis of corpus data, the zero occurrences in some subcorpora are worthy of careful consideration (Gablasova et al. 2017). For example, the absence of misuse of sentence-initial *Especially* falls into two categories, as shown in Table 9 below, the data of which are taken from Table 17 in section 4.2.

Table 9. Normalized Frequencies of Sentence-Initial *Especially*

Subcorpus	Misused Occurrences of Sentence-Initial *Especially*	Total Occurrences of Sentence-Initial *Especially*
CHN A2	0	176
IDN A2	0	136
TWN B1_2	0	70
CHN B2+	0	0
IDN B2+	0	0
THA B2+	0	0
TWN B2+	0	0

The first case (represented by the learners in subcorpora CHN A2, IDN A2, and TWN B1_2) may signal these writers' mastery of correct production of the target language feature (i.e., sentence-initial *Especially*) or their ability to correctly produce it. On the other hand, the second case (represented by the learners in the B2+ subcorpora of CHN, IDN, THA, and TWN) may signal the writers' choice of not producing sentence-initial *Especially* or their lack of an opportunity to produce it. This may be related to the small size of the relevant subcorpora: CHN B2+ (6,561 tokens; 13 participants), IDN B2+ (1,541 tokens; 3 participants), THA B2+ (1,016 tokens; 2 participants), and TWN B2+ (11,014 tokens; 23 participants). Thus, when answering the third research question (i.e., the question regarding the effect of L1 background and L2 proficiency on the frequency of misused sentence-initial *Especially*), all of the B2+ subcorpora were excluded from statistical analyses. Instead, statistical analyses were based on the tripartite proficiency-based classification of EFL learners: A2, B1_1, and B1_2.

To answer the third research question, R software version 4.3.1 (R Core Team 2023) was used for statistical

analyses, as well as for data visualization. The independent variables are L1 background (Korea, China, Indonesia, Japan, Thailand, and Taiwan) and level of English proficiency (A2, B1_1, and B1_2). The dependent variable is the frequency of misused sentence-initial *Especially*. For each of these variables, a 6 x 3 contingency table was analyzed using Pearson's chi-squared test. In corpus linguistics, chi-squared tests are commonly used to analyze categorical variables. Researchers, however, have pointed out that this traditional approach is not without problems (Bestgen 2014, Gablasova et al. 2017, Gries 2015). The underlying assumption of chi-squared tests is that "the data points are independent of each other" (Gries 2015, p. 101). Yet this assumption does not hold in corpus research, as explained in the following quote:

> [I]n a corpus of one million words, the chi-squared test works with the assumption that we have one million participants who are independent of each other. When we find that the target word or phrase occurred 5,000 times in the corpus, the test works with the information that 5,000 of your participants answered "yes" to the question while another 995,000 said "no." Language, however, does not work like this because individual words are not independent; on the contrary, individual words form an interconnected chain of meaning relationships as they combine in sentences and texts.
>
> (Gablasova et al. 2017, p. 150)

As these researchers suggested, analysis of variance (ANOVA) or its nonparametric equivalents can be to some extent complementary (Bestgen 2014, Gablasova et al. 2017, Gilquin and Granger 2015, Gries 2015). Taking this into consideration, I also conducted the Scheirer-Ray-Hare test (Scheirer et al. 1976, Sokal and Rohlf 1995), a nonparametric counterpart of two-way ANOVA, followed by a post-hoc test (the Dunn test, Dunn 1964). To this end, I calculated mean normalized frequencies of misused sentence-initial *Especially*. The data in Table 10 illustrate why it is essential to use mean rather than median. In the subcorpus of THA A2, the median normalized frequency is 19, while the mean normalized frequency is 23. The median, therefore, does not represent interwriter variation.

Table 10. An Example Spreadsheet of Misused Sentence-Initial *Especially*: The Case of THA A2 Subcorpus

L1 Background	L2 Proficiency	Code	RF	NF
THA	A2	294	2	38
THA	A2	190	1	19
THA	A2	230	1	19
THA	A2	344	1	19
THA	A2	364	1	19

Note. RF: raw frequency, NF: normalized frequency.

In the EFL subcorpora of the ICNALE Written Essays, three learners (KOR B2+ 219, JPN A2 109, and THA A2 294) produced misused sentence-initial *Especially* twice. All the other learners produced it only once. In corpus research, chi-squared tests are carried out on the total frequency of a certain expression in each subcorpus. For example, the total normalized frequency of misused sentence-initial *Especially* in THA A2 subcorpus is 114 (see Table 10 above and Table 19 in section 4.3.1). As shown in Table 10, the Scheirer-Ray-Hare test, as well as two-way ANOVA, takes individual texts/writers as observations. Because the sample size for each group (i.e., subcorpus) was small, the nonparametric version was used. The effect sizes were also measured in the analyses and reported as Cramer's *V* for Pearson's chi-squared test and the epsilon squared (ε^2) for the nonparametric ANOVA.

4. Results

4.1 *Especially* and Two Other Synonymous Particularizers in L1 and L2 Writing

The first research question focused on the distribution of *especially*, *particularly*, and *in particular* in L1 and L2 writing, and investigated their overuse and underuse patterns. The rationale for selecting these three particularizers was explained in section 2.[4] Table 11 and Figure 5 provide their frequency distribution in sentence-initial and non-sentence-initial positions of L1 and L2 writing.

Table 11. Raw and Normalized Frequencies of Three Particularizers in the ICNALE Written Essays Subcorpora

	Especially				*Particularly*				*In Particular*			
	Non-Initial		Initial		Non-Initial		Initial		Non-Initial		Initial	
Subcorpus	RF	**NF**	RF	**NF**	RF	**NF**	RF	**NF**	RF	**NF**	RF	**NF**
ENS students	26	**572**	1	**22**	5	**110**	0	**0**	2	**44**	0	**0**
ENS teachers	5	**250**	1	**50**	4	**200**	1	**50**	3	**150**	0	**0**
ENS S USA	24	**672**	1	**28**	3	**84**	0	**0**	2	**56**	0	**0**
ENS S GBR	0	**0**	0	**0**	0	**0**	0	**0**	0	**0**	0	**0**
ENS T USA	0	**0**	0	**0**	2	**360**	0	**0**	0	**0**	0	**0**
ENS T GBR	4	**716**	0	**0**	2	**358**	1	**179**	0	**0**	0	**0**
KOR A2	16	**480**	12	**360**	0	**0**	0	**0**	0	**0**	0	**0**
KOR B1_1	8	**296**	7	**259**	0	**0**	1	**37**	0	**0**	1	**37**
KOR B1_2	22	**550**	14	**350**	0	**0**	2	**50**	0	**0**	1	**25**
KOR B2+	19	**532**	12	**336**	2	**56**	0	**0**	0	**0**	4	**112**
CHN A2	14	**616**	4	**176**	0	**0**	0	**0**	0	**0**	0	**0**
CHN B1_1	104	**936**	10	**90**	1	**9**	1	**9**	1	**9**	0	**0**
CHN B1_2	50	**950**	8	**152**	2	**38**	0	**0**	0	**0**	0	**0**
CHN B2+	7	**1,064**	0	**0**	0	**0**	0	**0**	0	**0**	0	**0**
IDN A2	25	**1,700**	2	**136**	0	**0**	0	**0**	0	**0**	0	**0**
IDN B1_1	38	**1,026**	4	**108**	1	**27**	0	**0**	2	**54**	0	**0**
IDN B1_2	40	**1,000**	3	**75**	1	**25**	0	**0**	0	**0**	0	**0**
IDN B2+	3	**1,947**	0	**0**	0	**0**	0	**0**	0	**0**	0	**0**
JPN A2	22	**330**	15	**225**	3	**45**	2	**30**	2	**30**	3	**45**
JPN B1_1	32	**416**	23	**299**	1	**13**	2	**26**	0	**0**	0	**0**
JPN B1_2	7	**315**	4	**180**	0	**0**	0	**0**	0	**0**	0	**0**
JPN B2+	3	**351**	3	**351**	1	**117**	0	**0**	0	**0**	0	**0**
THA A2	16	**304**	7	**133**	2	**38**	0	**0**	1	**19**	1	**19**
THA B1_1	56	**672**	21	**252**	3	**36**	0	**0**	0	**0**	0	**0**
THA B1_2	31	**682**	10	**220**	0	**0**	0	**0**	1	**22**	0	**0**
THA B2+	0	**0**	0	**0**	0	**0**	0	**0**	0	**0**	0	**0**
TWN A2	7	**546**	2	**156**	0	**0**	0	**0**	0	**0**	0	**0**
TWN B1_1	14	**350**	7	**175**	0	**0**	1	**25**	1	**25**	0	**0**
TWN B1_2	18	**630**	2	**70**	0	**0**	0	**0**	0	**0**	0	**0**
TWN B2+	7	**637**	0	**0**	0	**0**	0	**0**	0	**0**	0	**0**

Note. (i) S: students, T: teachers, RF: raw frequency, NF: normalized frequency, Non-Initial: non-sentence-initial position, Initial: sentence-initial position. (ii) NF = RF × the number of one occurrence per million words in each subcorpus (shown in the last column of Table 7).

[4] No overuse/underuse patterns were observed for the three other particularizers in Table 4.

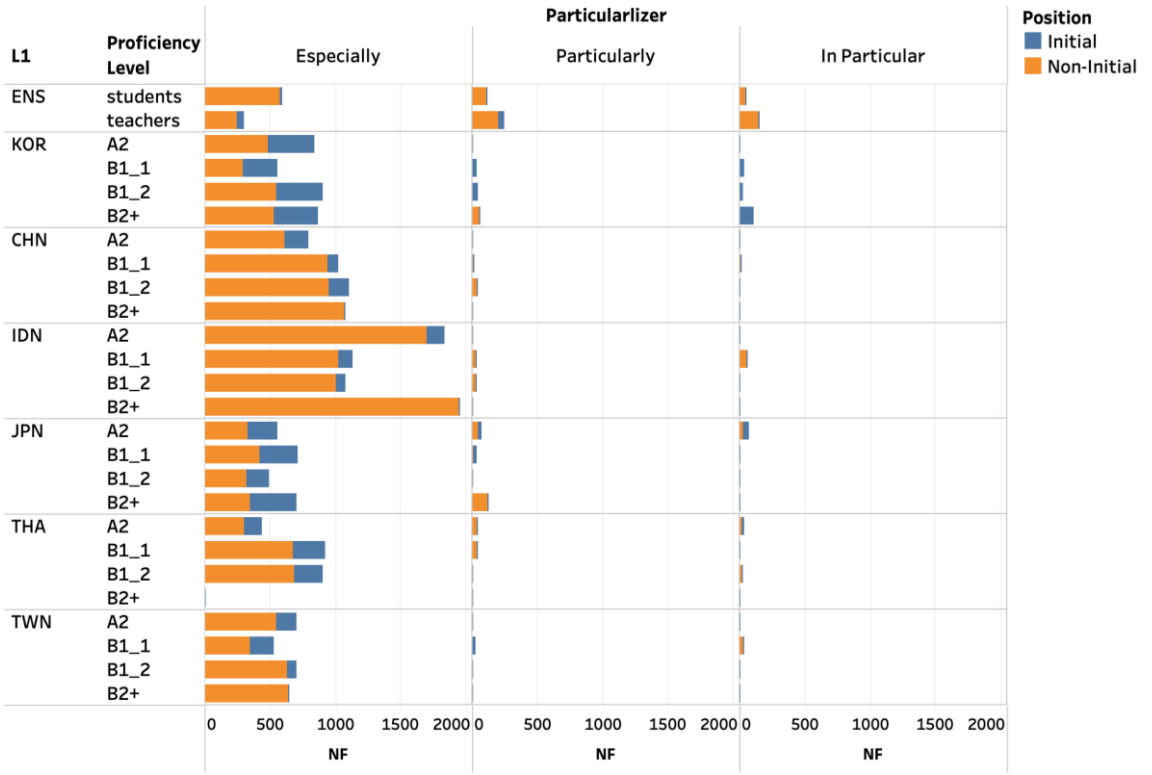

Figure 5. Normalized Frequencies of Three Particularizers in the ICNALE Written Essays Subcorpora

Keyword analyses were conducted by specifying one of the 24 EFL subcorpora in Table 11 as the target and the subcorpus of ENS students as the reference. In this section, particularizers overused or underused in L2 writing in comparison with those in L1 writing are grouped according to their type and arranged in the descending order of the chi-squared score. In Tables 12-16, a high chi-squared value (x^2) suggests that the particularizer under consideration appears extraordinarily more or less in an EFL subcorpus (i.e., the target corpus) than in the subcorpus of ENS students (i.e., the reference corpus). Table 12 presents the list of subcorpora in which sentence-initial *Especially* is overused. As shown in Table 12, Korean and Japanese learners at all four proficiency levels overuse sentence-initial *Especially*. The top three highest scores are obtained by Korean learners (A2, B1_2, and B2+). Sentence-initial *Especially* is also overused by three levels of Thai and Chinese learners (A2, B1_1, and B1_2), two levels of Taiwanese learners (A2 and B1_1) and one level of Indonesian learners (B1_2). The degree of overuse varies greatly across these subcorpora.

Table 12. Overuse of Sentence-Initial *Especially*

Subcorpus	x^2	Subcorpus	x^2	Subcorpus	x^2
KOR A2	13.23	KOR B1_1	8.55	CHN B1_2	4.46
KOR B1_2	12.84	JPN A2	7.3	THA A2	3.24
KOR B2+	11.84	THA_B1_2	6.91	TWN A2	1.93
JPN B1_1	10.53	TWN B1_1	5.01	CHN B1_1	1.76
JPN B2+	9.41	JPN B1_2	4.52	IDN B1_2	1
THA B1_1	9.29	CHN A2	4.51		

Table 13 shows the list of subcorpora in which non-sentence-initial *especially* (marked by a small letter) is overused. Indonesian learners at all four proficiency levels overuse non-sentence-initial *especially*. This particularizer is also overused by three levels of Chinese learners (B1_1, B1_2, and B2+), two levels of Thai learners (B1_1 and B1_2), and one level of Taiwanese learners (B2+). In the case of Thai and Taiwanese learners, the chi-squared values are extremely low (less than one).

Table 13. Overuse of Non-Sentence-Initial *Especially*

Subcorpus	x^2	Subcorpus	x^2	Subcorpus	x^2
IDN A2	15.97	CHN B1_2	4.6	THA B1_2	0.27
CHN B1_1	5.17	IDN B2+	4.3	TWN B2+	0.08
IDN B1_2	4.92	CHN B2+	2.19		
IDN B1_1	4.84	THA B1_1	0.54		

The underuse of the three particularizers is also noticeable in non-sentence-initial positions. As shown in Table 14, Thai A2 learners and Japanese learners at all four proficiency levels underuse non-sentence-initial *especially*. Korean learners at B1_1 level and those at A2 and B2+ levels also underuse it although in the latter case, the chi-squared scores are extremely low (less than one). Table 15 reveals that non-sentence-initial *particularly* is underused by Chinese A2 level learners (x^2 = 8.09), Japanese B1_1 level learners (x^2 = 5.51) and some other learner groups. As shown in Table 16, the expression *particular* in non-sentence-initial position is underused by Thai B1_1 level learners (x^2 = 3.48) and several other learner groups. However, care should be taken in interpreting the results in Table 16. The keyword analysis identifies individual words rather than multiword units (e.g., fixed expressions such as *in particular*). Therefore, the results in Table 16 contain instances of *particular* used as an adjective (as in *particular dish*).

Table 14. Underuse of Non-Sentence-Initial *Especially*

Subcorpus	x^2	Subcorpus	x^2	Subcorpus	x^2
THA A2	4.46	JPN B1_2	2.04	KOR A2	0.28
JPN A2	4.13	JPN B1_1	1.86	KOR B2+	0.04
KOR B1_1	2.63	JPN B2+	0.51		

Table 15. Underuse of Non-Sentence-Initial *Particularly*

Subcorpus	x^2	Subcorpus	x^2	Subcorpus	x^2
CHN A2	8.09	IDN B1_2	2	JPN A2	1.56
JPN B1_1	5.51	CHN B1_2	1.72	KOR B2+	0.57
THA B1_1	2.16	THA A2	1.72		

Table 16. Underuse of Non-Sentence-Initial *Particular*

Subcorpus	x^2	Subcorpus	x^2	Subcorpus	x^2
THA B1_1	3.48	THA B1_2	1.59	TWN B1_1	0.79
CHN B1_2	1.6	KOR B1_2	0.79	THA A2	0.66

The results of the keyword analysis conducted between the whole EFL learners as the target and ENS students as the reference indicated that there was a strong tendency for EFL learners to overuse sentence-initial *Especially* (x^2 = 6.76) while there was only a very weak tendency to overuse non-sentence-initial *especially* (x^2 = 0.25). They also showed a very strong tendency to underuse non-sentence-initial *particularly* (x^2 = 14.72) and a

tendency to underuse non-sentence-initial *particular* (x^2 = 3.16). As noted earlier, the latter requires a cautious interpretation of the results.

4.2 The Uses of Sentence-Initial *Especially* in L1 and L2 Writing

The second research question examined the uses of sentence-initial *Especially* in L1 and L2 writing. As shown in Table 11 and (13), only one token was found in the subcorpus of ENS students.

(13) ENS students

Especially in the U. S., most of the food is already so bad for you that having the extra affect [*sic*] if [*sic*] smoking is almost just like slapping yourself in the face.

(WE_ENS_SMK0_061_XX_1)

In (13), *the extra affect if smoking* should be read as *the extra effect of smoking*. Example (13) was produced by an ENS participant from the USA. In (13), a prepositional phrase (PP) follows sentence-initial *Especially*, as in the first example in Table 3. There was also only one token of sentence-initial *Especially* in the subcorpus of ENS teachers, as shown in Table 11 and (14).

(14) ENS teachers

I agree with this statement. ***Especially*** for those students who do not have to pay for their education or day to day living expenses. In my experience, I find that these students in particular, have the lowest rate of successfully gaining employment within the first three months of entering the job market.

(WE_ENS_PTJ0_135_XX_2)

Example (14) was written by an ENS participant from New Zealand. Although the sentence-initial *Especially* in (14) is also immediately followed by a PP, unlike (13), (14) is not a full sentence but a sentence fragment. In Table 3, examples like (14) are labeled as ‘sentence fragment (PP)’.

The uses of sentence-initial *Especially* in L2 writing are categorized in Table 17 and Figure 6 according to the constituents that follow it. In Table 17 and Figure 6, the category ‘the subject of a clause’ refers to examples such as those presented in (15).

(15) a. Especially, a restaurant is the place where there are many kids.

(WE_KOR_SMK0_064_A2_0)

b. Especially I hate people who smoke during eating.

(WE_JPN_SMK0_377_B1_1)

c. Especially the smell of smoke reduces the taste of other people’s food.

(WE_KOR_SMK0_276_B1_2)

d. Especially, it is a big burden for both students and parents to make the tuition fee.

(WE_KOR_PTJ0_219_B2_0)

Table 17. Normalized Frequencies of the Constituents That Follow Sentence-Initial *Especially* in L2 Writing

	The Subject of a Clause	PP	Adverb	*If/When*-Clause	Sentence Fragment	Other	Totals
KOR A2	240	0	30	30	60	0	360
KOR B1_1	148	37	0	37	37	0	259
KOR B1_2	200	75	0	25	50	0	350
KOR B2+	224	84	0	0	28	0	336
CHN A2	0	176	0	0	0	0	176
CHN B1_1	18	36	9	9	9	9	90
CHN B1_2	19	114	0	0	19	0	152
CHN B2+	0	0	0	0	0	0	0
IDN A2	0	0	0	0	68	68	136
IDN B1_1	27	0	0	0	81	0	108
IDN B1_2	25	0	0	0	50	0	75
IDN B2+	0	0	0	0	0	0	0
JPN A2	165	60	0	0	0	0	225
JPN B1_1	208	26	0	26	26	13	299
JPN B1_2	90	45	0	0	45	0	180
JPN B2+	117	117	0	117	0	0	351
THA A2	114	0	0	0	19	0	133
THA B1_1	24	24	0	12	144	48	252
THA B1_2	44	66	0	0	44	66	220
THA B2+	0	0	0	0	0	0	0
TWN A2	78	0	0	0	78	0	156
TWN B1_1	75	75	0	0	25	0	175
TWN B1_2	0	35	0	0	0	35	70
TWN B2+	0	0	0	0	0	0	0
Totals	1,816	970	39	256	783	239	4,103

Note. PP = prepositional phrase.

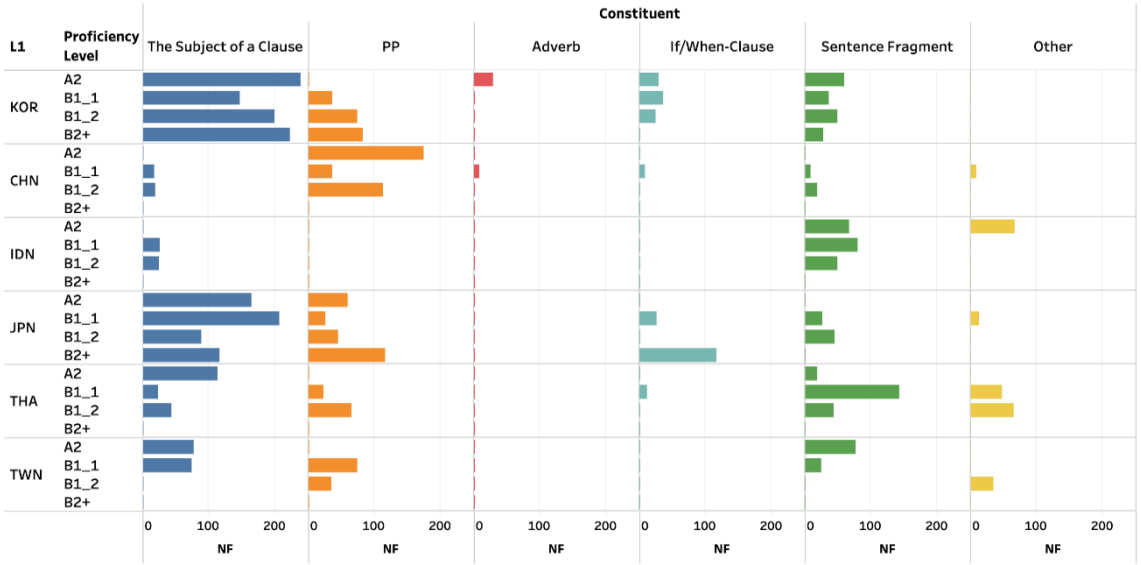

Figure 6. Normalized Frequencies of the Constituents That Follow Sentence-Initial *Especially* in L2 Writing

Throughout this paper, this sentence-initial use of *Especially* is treated as misuse. Constituents such as PP, adverb, *if*/*when*-clause, and sentence fragment have already been noted in Table 3. The category 'other' refers to cases that do not fall into any of the categories described in Table 3, and includes the following instances:

(16) a. Especially pregnant women's case, smoking affects unborn babies.
(WE_JPN_SMK0_320_B1_1)

b. Especially, what a bad impression it will leave to visitors from other cities or even other countries!
(WE_CHN_SMK0_119_B1_1)

c. Especially smoking, it should be completely banned at all the restaurants in the country.
(WE_THA_SMK0_248_B1_1)

In (16a), the preposition *in* is missing before the noun phrase *pregnant women's case*. In (16b), sentence-initial *Especially* is followed by an exclamative clause whereas in (16c), it is followed by a left dislocation construction. Examples like (16c) were produced by Thai B1_1 level learners only, and there were four such tokens in this subcorpus. In the present study, a left dislocated noun phrase is distinguished from the subject of a clause. Further discussion is beyond the scope of this paper. Instances of sentence-initial *Especially* that belong to the category 'other' are negligible (5.83%, see Table 18) and therefore will not be of further concern in this study.

The results in Table 17 and Figure 6 are summarized in Table 18 and Figure 7. As confirmed in the previous section, there is a strong tendency for EFL learners to overuse sentence-initial *Especially*. The results in this section reveal that almost half of these overused instances (44.26%) are cases where sentence-initial *Especially* is followed by the subject of a clause. As can be seen from Table 18, the proportion of misuse (errors) over the total instances of sentence-initial *Especially* varies considerably depending on L1 background: Korean (62.22%), Japanese (54.98%), Taiwanese (38.15%), Thai (30.08%), Indonesian (16.30%), and Chinese (8.85%). The impact of L1 background on misused sentence-initial *Especially* will be discussed in more depth in the next section.

Table 18. Summary of Table 17 Across L1 Backgrounds

	The Subject of a Clause	PP	Adverb	*If*/*When*-Clause	Sentence Fragment	Other	Totals
KOR	812	196	30	92	175	0	1,305
	(62.22)	(15.02)	(2.30)	(7.05)	(13.41)	(0.00)	(100)
CHN	37	326	9	9	28	9	418
	(8.85)	(77.99)	(2.15)	(2.15)	(6.70)	(2.15)	(100)
IDN	52	0	0	0	199	68	319
	(16.30)	(0.00)	(0.00)	(0.00)	(62.38)	(21.32)	(100)
JPN	580	248	0	143	71	13	1,055
	(54.98)	(23.51)	(0.00)	(13.55)	(6.73)	(1.23)	(100)
THA	182	90	0	12	207	114	605
	(30.08)	(14.88)	(0.00)	(1.98)	(34.21)	(18.84)	(100)
TWN	153	110	0	0	103	35	401
	(38.15)	(27.43)	(0.00)	(0.00)	(25.69)	(8.73)	(100)
Totals	**1,816**	**970**	**39**	**256**	**783**	**239**	**4,103**
	(44.26)	**(23.64)**	**(0.95)**	**(6.24)**	**(19.08)**	**(5.83)**	**(100)**

Note. Percentages are presented in parentheses. PP = prepositional phrase.

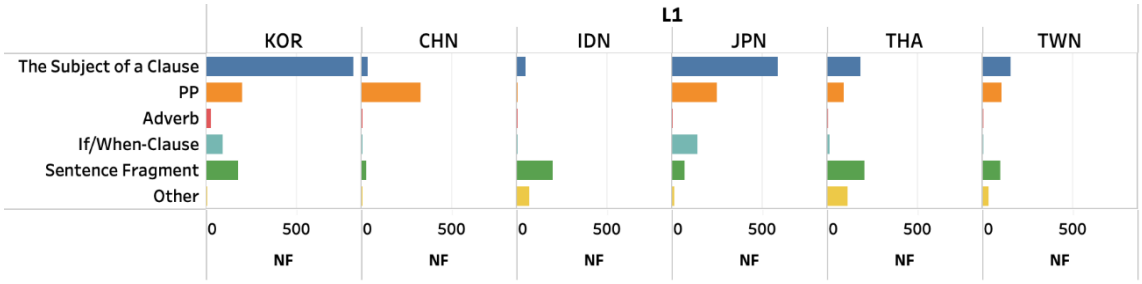

Figure 7. Summary of Figure 6 Across L1 Backgrounds

4.3 The Effect of L1 Background and L2 Proficiency on the Frequency of Misused Sentence-Initial *Especially*

The third research question aimed to find out whether L1 background and L2 proficiency affect the frequency of misused sentence-initial *Especially* in L2 writing. Of the various instances of sentence-initial *Especially*, described in Tables 17 and 18, misuse refers to cases where sentence-initial *Especially* precedes the subject of a clause. As explained in section 3.3, learners at three proficiency levels (A2, B1_1, and B1_2) are considered in this section.

4.3.1 Results of Pearson's Chi-Squared Test

First, Table 19 displays normalized frequencies of misused sentence-initial *Especially* across L1 backgrounds and the three CEFR levels. Pearson's chi-squared test was performed to assess the association between L1 background and English proficiency. The relationship between these variables was statistically significant (x^2 = 217.70, *df* = 10, *p*<.001). The effect size was moderate with Cramer's *V* = .272. Overall, the highest frequency of misused sentence-initial *Especially* was found in the L1 Korean subcorpus (39.86%), followed by Japanese (31.39%), Thai (12.34%), Taiwanese (10.37%), Indonesian (3.53%), and Chinese (2.51%). As shown in Table 19, the total frequency is 100.

Table 19. Normalized Frequencies of Misused Sentence-Initial *Especially* Across L1 Backgrounds and the Three CEFR L2 Proficiency Levels

	KOR	CHN	IDN	JPN	THA	TWN	Totals
A2	240	0	0	165	114	78	597
	(16.27)	(0.00)	(0.00)	(11.19)	(7.73)	(5.29)	(40.47)
B1_1	148	18	27	208	24	75	500
	(10.03)	(1.22)	(1.83)	(14.10)	(1.63)	(5.08)	(33.90)
B1_2	200	19	25	90	44	0	378
	(13.56)	(1.29)	(1.69)	(6.10)	(2.98)	(0.00)	(25.63)
Totals	588	37	52	463	182	153	1,475
	(39.86)	(2.51)	(3.53)	(31.39)	(12.34)	(10.37)	(100)

Note. Percentages are presented in parentheses.

L1 Korean learners at A2 level exhibited the highest frequency (16.27%) of all 18 subcorpora. Of the three proficiency-based subcorpora involving L1 Korean learners, those at B1_1 level had the lowest frequency (10.03%) while an increasing tendency was observed at B1_2 level (13.56%). In the case of Chinese and Indonesian learners,

the frequency was extremely low at A2 level (0%). The frequency was similar for learners at B1_1 and B1_2 levels although it was very low (around 2%). As for Japanese learners, the frequency increased from 11.19% at A2 level to 14.10% at B1_1 level, but it was lowest at B1_2 level (6.10%). Thai learners showed the highest frequency at A2 level (7.73%). At the levels of B1_1 and B1_2, the frequency was low (approximately 3%). Taiwanese learners show a similar level of frequency at the levels of A2 (5.29%) and B1_1 (5.08%), but in the case of B1_2 level, the frequency of the misuse of sentence-initial *Especially* converges to zero.

4.3.2 Results of the Scheirer-Ray-Hare Test and the Post-Hoc Dunn's Test

Table 20 presents mean normalized frequencies of misused sentence-initial *Especially* across L1 backgrounds and the three CEFR levels. Figure 8 shows the interaction plot from the data in Table 20.

Table 20. Mean Normalized Frequencies of Misused Sentence-Initial *Especially* Across L1 Backgrounds and the Three CEFR L2 Proficiency Levels

	KOR	CHN	IDN	JPN	THA	TWN
A2	30	0	0	17	23	78
B1_1	37	9	27	13	12	25
B1_2	25	19	25	45	22	0

Note. All standard deviations are zero except for: (i) JPN A2 = 10.6, (ii) THA A2 = 13.4.

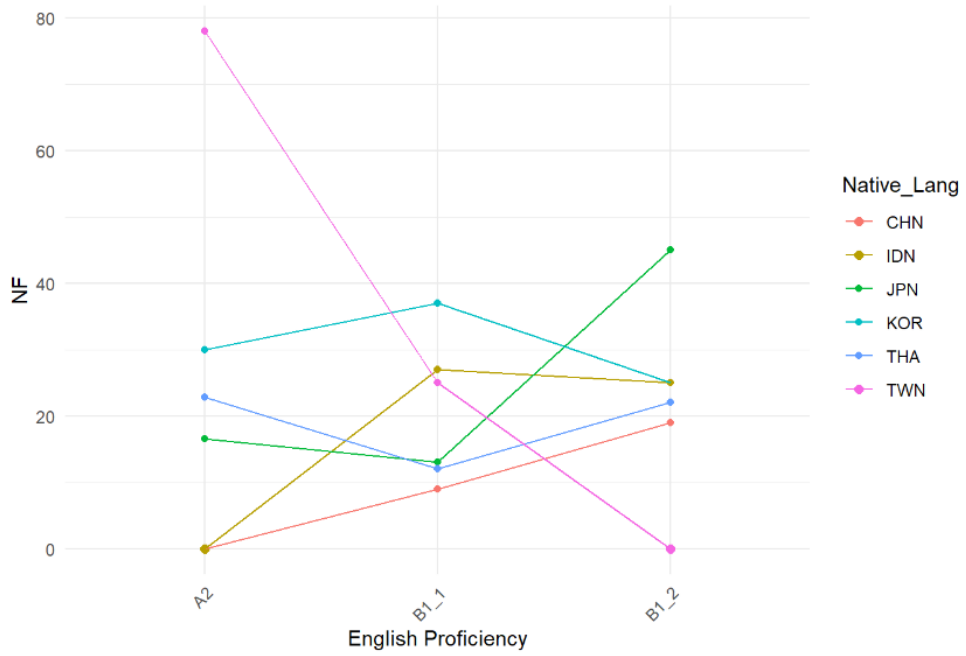

Figure 8. Interaction Plot of Mean Normalized Frequencies (NFs) of Misused Sentence-Initial *Especially* by Native Language and English Proficiency

The Scheirer-Ray-Hare test was conducted to analyze the effect of L1 background and English proficiency on the frequency of misused sentence-initial *Especially*. Table 21 shows Scheirer-Ray-Hare test results and effect size. There was a significant interaction between L1 background and English proficiency (H = 17.86, df = 7, p = .013). Despite reaching statistical significance, the effect size was very small (ε^2 = .018). Simple main effects analysis showed that L1 background did have a statistically significant effect on the frequency of misused sentence-initial *Especially* (H = 26.03, df = 5, p = .000). The effect size was large (ε^2 = .350). Simple main effects analysis showed that English proficiency had a statistically significant effect on the frequency of misused sentence-initial *Especially* (H = 7.96, df = 2, p = .019). The effect size was small (ε^2 = .095).

Table 21. Scheirer-Ray-Hare Test Results and Effect Size

Source	*df*	Sum of Squares (SS)	*H*	*p*	ε^2
Native Language (N)	5	9343.1	26.03***	.000	.350
English Proficiency (E)	2	2858.3	7.96*	.019	.095
N × E	7	6409.8	17.86*	.013	.018
Residuals	51	1525.7			

*p<.05, **p<.01, ***p<.001

In order to find out whether all six of the L1s are different from each other, the Dunn test was conducted, and the results are shown in Table 22. Korean learners differed significantly from Chinese (p < .05) and Japanese (p < .001) learners. All the other pairwise comparisons were not significant.

Table 22. Post-Hoc Analysis of the Frequency of Misused Sentence-Initial *Especially* by L1 Background

Pairwise Comparisons	*Z*	Significance
KOR vs. CHN	3.33	*
KOR vs. IDN	0.35	
KOR vs. JPN	5.23	***
KOR vs. THA	2.74	
KOR vs. TWN	0.22	
CHN vs. IDN	-1.97	
CHN vs. JPN	-0.87	
CHN vs. THA	-1.44	
CHN vs. TWN	-2.54	
IDN vs. JPN	1.73	
IDN vs. THA	1.07	
IDN vs. TWN	-0.16	
JPN vs. THA	-1.13	
JPN vs. TWN	-2.63	
THA vs. TWN	-1.62	

*p<.05, **p<.01, ***p<.001

Table 23 presents the results of post-hoc comparisons of three L2 proficiency levels. A2 level learners produced misused sentence-initial *Especially* more frequently than B1_1 level learners (p < .01), who in turn produced it less frequently than B1_2 level learners (p < .01). The latter finding is contrary to the common belief that L2 learners' error rates generally decrease as their proficiency level increases. A2 level and B1_2 level learners were not significantly different from each other.

Table 23. Post-Hoc Analysis of the Frequency of Misused Sentence-Initial *Especially* by English Proficiency

Pairwise Comparisons	*Z*	Significance
A2 vs. B1_1	3.39	**
A2 vs. B1_2	-0.56	
B1_1 vs. B1_2	-3.46	**

*p<.05, **p<.01, ***p<.001

In Table 24, only the comparisons with significant interaction effects are presented instead of all possible comparisons. Korean B1_1 level learners produced misused sentence-initial *Especially* more frequently than the same level learners from the following L1s: Japanese (p < .001) and Chinese (p < .05). Japanese B1_1 level learners produced it less frequently than Korean A2 level learners (p < .001) and Korean B1_2 level learners (p < .05). They also produced it less frequently than Japanese B1_2 level learners (p < .05). Finally, Chinese B1_1 level learners produced misused sentence-initial *Especially* less frequently than Korean A2 level learners (p < .05).

Table 24. Post-Hoc Analysis of the Frequency of Misused Sentence-Initial *Especially* by L1 Background and English Proficiency

Pairwise Comparisons	*Z*	Significance
KOR B1_1 vs. JPN B1_1	4.53	***
KOR B1_1 vs. CHN B1_1	3.60	*
JPN B1_1 vs. KOR A2	-5.06	***
JPN B1_1 vs. KOR B1_2	-3.66	*
JPN B1_1 vs. JPN B1_2	-3.66	*
CHN B1_1 vs. KOR A2	-3.51	*

*p<.05, **p<.01, ***p<.001

In summary, the comparison and contrast of the frequency of misused sentence-initial *Especially* under the two different conditions, L1 background and English proficiency, show that L1 background indeed has a significant effect with a large effect size. Specifically, the results of the present study showed a relatively high frequency of misused sentence-initial *Especially* in the writings of Korean learners (as opposed to Japanese and Chinese learners). In addition, there is a statistically significant association between English proficiency and the frequency of misused sentence-initial *Especially* in EFL learner writing. The effect size, however, is small, suggesting a limited impact. The prevalence of misuse among A2 level learners (compared to B1_1 level learners) is not unexpected. Interestingly, however, it was found that B1_2 level learners produced misused sentence-initial *Especially* more frequently than B1_1 level learners. This increasing tendency of errors is not consistent with the general interlanguage development path of L2 learners. As for the interaction between L1 background and English proficiency, the effect size is very small (the Scheirer-Ray-Hare test) or moderate (Pearson's chi-squared test) although the results are statistically significant in both tests. This indicates that the strength of the correlation between the two independent variables is not strong.

5. Discussion

The present study aimed to determine whether misused sentence-initial *Especially* is an L1-related production tendency or a universal feature of interlanguage development. The most crucial finding was that L1 background

did have a statistically significant effect on the frequency of misused sentence-initial *Especially*. It is also notable that the results yielded a large effect size. The relatively high frequency of sentence-initial *Especially* by L1 Korean writers can be attributed to L1 transfer. In an attempt to understand why Korean EFL learners misuse sentence-initial *Especially* more frequently than other Asian EFL learners, this study conducted a follow-up investigation of a Korean language corpus. Specifically, the current study searched the National Institute of Korean Language (NIKL) Written Corpus (version 1.2, National Institute of Korean Language 2020). The NIKL Written Corpus is a collection of 10,045 JavaScript Object Notation (JSON) files, and includes samples of written language from a wide range of sources, excluding newspapers. In more detail, it consists of informational books (86.12%), imaginary books (13.87%), and magazines (0.01%). The books are further divided into two subcorpora: one constructed in the year 2018 and the other constructed in the year 2019. In this study, the 2018 informative book subcorpus (which comprises 107 JSON files) was chosen due to the comparability with other corpora (such as the ICNALE and the COCA academic subcorpus), as well as for reasons of time and efficiency. Its total number of tokens, calculated using AntConc (version 4.3.0), is 6,181,830. The KWIC tool in AntConc was used to search for instances of *thukhi* ('especially') in this subcorpus. As shown in Table 25, the proportional use of the Korean particularizer was 53% in initial position compared to 47% in medial position.

Table 25. Distribution of the Korean Equivalent of *Especially* in the NIKL 2018 Informative Book Subcorpus Across Positions

Position	Raw Frequency	Normalized Frequency	Percentage
Initial	2,570	416	53
Medial	2,320	375	47
Total	4,890	791	100

Note. Normalized frequency = occurrence per million words.

Regarding the distribution of the Korean particularizer *thukhi* ('especially') shown in Table 25, one may say that initial positioning occurs only slightly more frequently than medial positioning. However, given the overall low use of initial position in the COCA academic section (see Table 4), the frequent misuse of sentence-initial *Especially* among L1 Korean writers of L2 English can indeed be associated with a characteristic specific to this particular L1 background, namely the particularizer's preference for initial position in the Korean language. In short, L1 transfer is at least partially responsible for the high frequency of misused sentence-initial *Especially* by L1 Korean writers.

A second important finding from this study is that L2 proficiency plays a statistically significant role in the frequency of misused sentence-initial *Especially* in L2 writing. The small effect size indicates, however, that the correlation is restricted. A2 level learners produced misused sentence-initial *Especially* more frequently than B1_1 level learners, as expected. Yet, contrary to the common assumption that an increase of L2 input results in a decrease of L2 errors, B1_2 level learners produced misused sentence-initial *Especially* more frequently than B1_1 level learners. A further problem arises when considering B2+ level learners, who were excluded from statistical analyses. As can be seen in Table 17, of the four proficiency-based subcorpora involving L1 Korean EFL learners, those at B2+ level showed the second highest frequency, while those at A2 level showed the highest frequency of misused sentence-initial *Especially*. Such an increasing tendency of errors is almost inconceivable if misused sentence-initial *Especially* is a universal feature of interlanguage development because intralingual or developmental errors usually disappear as L2 proficiency increases. On the contrary, the increasing misuse of sentence-initial *Especially* among L1 Korean EFL learners can be accounted for by the fact that unlike English,

Korean allows placing the subject of a clause immediately after the Korean particularizer *thukhi* (‘especially’), as in (17).[5]

(17) thukhi pwuchin-i hwulyunghan susungtul-ul mosiko wass-ta
특히 부친이 훌륭한 스승들을 모시고 왔다.
especially father-NOM great teachers-ACC brought-DECL
‘In particular, my father brought great teachers.’

(NIKL, WBRW1800000084)

Example (17), which is taken from the Korean language corpus examined in this study, is in stark contrast to (4b), (5b), (6b), (7c), and (8b). In sum, the second major finding from this study also lends support to an alternative account of the misuse phenomenon based on L1 transfer.

A third noteworthy finding is the infrequent use of non-sentence-initial *especially* (not to mention sentence-initial *Especially*) among ENS teachers (see Table 11). An important issue raised by Callies (2015) is the question of “against which yardstick learner data should be compared and evaluated” (p. 40). The present study followed the guidelines offered by Ishikawa (2023), who distinguished a peer reference from a pedagogical reference. Thus, quantitative differences such as overuse and underuse were calculated by considering the subcorpus of ENS students as a reference corpus. Nevertheless, it is worth exploring whether there remains a possibility that expert writers (i.e., ENS teachers), as opposed to novice writers (i.e., ENS students), resort to alternative devices instead of using particularizers. As shown in (18), expressions such as *most importantly*, as well as *wh*-cleft constructions, may serve as a focusing device akin to particularizers.

(18) ENS teachers
a. They can also learn other skills such as simple workplace dynamics and *most importantly*, they would also be able to use the part-time job on their resumes to show that they have some sort of real work experiences. (WE_ENS_PTJ0_184_XX_2)
b. I think *what is more important than having a part time job* is getting a good average and doing well with studies. (WE_ENS_PTJ0_129_XX_2)

(19) a. They can also learn other skills such as simple workplace dynamics and *in particular*, they would also be able to use the part-time job on their resumes to show that they have some sort of real work experiences.
b. I think *in particular* that it is more important to get a good average and do well with studies than to have a part time job.

The expression *most importantly* means ‘above all’, and (18a) can be paraphrased as (19a). In a *wh*-cleft construction, the focused element is new information, which appears at the end (i.e., after the copula). (18b) can be restated as (19b). In order to strengthen the arguments, further testing with a larger-size corpus of L1 English expert writing is recommended.

A final finding worth noting is that unlike sentence-initial *Especially*, sentence-initial *Particularly* and sentence-

[5] In (17), the following abbreviations are used: (i) ACC: accusative, (ii) DECL: declarative ending, and (iii) NOM: nominative.

initial *In particular* were not overused by EFL learners (both as a whole and in different L1 groups). All three English expressions can be translated into *thukhi* in Korean although their interchangeability varies depending on their position and neighboring constituent. These three near-synonyms also vary in terms of their frequency in the academic section of the COCA, as shown in Table 4. Davies and Gardner (2010) provided a list of the top 5,000 most frequently occurring words in contemporary American English, with rank 1 being the definite article *the*. *Especially* received a ranking of 525 and *particularly* received a ranking of 782. The word *particular* (not the compound particularizer *in particular*) was ranked 952nd. The answer to the question of what words are overused by L2 learners seems to be related to the frequency of the near-synonyms in the target language. That is, the more frequent the occurrence of a near-synonym in the target language, the more likely it is to be overused by L2 learners.

The findings of this study raise several pedagogical implications for the teaching of the near-synonyms under discussion. The key pedagogical implication arising from the limited role of L2 input (discussed above with regard to the second finding) is that the placement of particularizers is an area where formal L2 instruction is effective and efficient. By means of data-driven learning (DDL, Johns 1991) activities, L2 learners can be encouraged to observe the positional patterns of these particularizers, thereby uncovering the constraint against placing sentence-initial *Especially* before the subject of a clause. Simply put, DDL or "discovery learning" (McEnery and Xiao 2011, p. 370) refers to the use of concordances in language teaching (Johns 1991). Concordances may be presented in a variety of ways, ranging from edited and truncated versions to original full sentences, in accordance with L2 learners' proficiency levels. As is well-known, DDL promotes learner autonomy (Boulton 2010). Admittedly, there are also some disadvantages of discovery learning, such as more preparation time needed for teachers, as well as more time needed for students to complete DDL activities (Lin and Lee 2015). An alternative way of self-guided learning is encouraging L2 learners to use corpus-based monolingual dictionaries, some of which include crucial information on the placement of particularizers. Depending on the learners' learning styles, providing explicit corrective feedback, such as explicit correction, would also be effective for remedying the misuse of sentence-initial *Especially*.

There are some limitations of the present study. First and foremost, the L1 transfer effect needs to be tested more thoroughly in future research. As reported at the end of section 4.1, the keyword analysis, the main purpose of which was to identify overuse/underuse patterns, revealed that compared to ENS students, EFL learners as a whole exhibited a strong tendency to overuse sentence-initial *Especially* (x^2 = 6.76). All of the Korean learners, regardless of their English proficiency levels, showed an extremely strong tendency toward this overuse pattern. Although their chi-squared values were relatively low compared to Korean learners, Japanese learners at A2, B1_1, and B2+ levels, as well as Thai learners at B1_1 and B1_2 levels also had a strong tendency to overuse sentence-initial *Especially* (x^2 > 6.76, see Table 12 for details). Yet, only the Korean L1 corpus was examined to explore the possibility of L1 influence. This calls for international research collaboration involving researchers from different L1 backgrounds. According to Zhang (2000), Chinese EFL learners' overuse of the initial positioning of conjunctions (*however*, *therefore*, etc.) is due to L1 transfer: the corresponding Chinese expressions are, for the most part, placed at the beginning of a sentence. With regard to the adverb *also*, Liu and Rhee (2017) found that Korean EFL learners had a higher proportional use of initial position than Chinese EFL learners, who showed a preference for pre-verbal position. Liu and Rhee (2017) also noted that Chinese EFL learners preferred to place adverbs such as *just* and *still* in post-verbal position. The post-hoc analysis, reported in section 4.3.2 in response to the third research question on the misuse (rather than overuse) of sentence-initial *Especially*, revealed that L1 Korean EFL learners differed significantly from L1 Chinese EFL learners (p < .05). Taken together, these considerations point to the need for examining the placement of particularizers in L1 Chinese writing and other L1 corpora. Future studies should also investigate multi-L1 learner corpora containing EFL learners from

European countries. Finally, it might also be beneficial in future research to compare American and British English in terms of the placement of particularizers.

6. Conclusion

Adverbial placement is a well-trodden topic in L2 acquisition, as well as corpus linguistics. However, very few studies have addressed the prohibition against placing the particularizing adverb *especially* at the beginning of a sentence before the subject. The current study investigated this issue by employing six subcorpora of the ICNALE (Korean, Chinese, Indonesian, Japanese, Taiwanese, and Thai EFL learners). It also attempted to contribute to the body of research concerning the role of L1 background and L2 proficiency on adverbial positions in L2 writing.

The findings of this study are summarized as follows. In response to the first research question, the present study found that compared to ENS students, EFL learners as a whole exhibited a strong tendency to overuse sentence-initial *Especially* in their writing of argumentative essays. All of the L1 Korean learners, regardless of their English proficiency levels, showed an extremely strong tendency toward this overuse pattern. As for the second research question, nearly half (44.26%) of the overused instances of sentence-initial *Especially* in L2 writing are cases where sentence-initial *Especially* is followed by the subject of a clause. The misuse of sentence-initial *Especially* is undoubtedly an interlanguage phenomenon. In the subcorpus of ENS students, there is only one token of sentence-initial *Especially*, which is followed by a prepositional phrase. The proportion of misuse over the total instances of sentence-initial *Especially* varies considerably depending on L1 background: Korean (62.22%), Japanese (54.98%), Taiwanese (38.15%), Thai (30.08%), Indonesian (16.30%), and Chinese (8.85%). In addressing the third research question, L1 background did have a statistically significant impact on the frequency of misused sentence-initial *Especially* in EFL learner writing. The effect size was large. In addition, there was a statistically significant association between English proficiency and the frequency of misused sentence-initial *Especially*. The effect size, however, was small, suggesting a limited impact. It is interesting to note that B1_2 level learners produced misused sentence-initial *Especially* more frequently than B1_1 level learners, which is a virtual impossibility if the misuse is considered to be an intralingual or developmental error.

The L1 transfer effect was demonstrated by a follow-up investigation of a Korean language corpus. Although the dataset was small, it was compiled from a large corpus, thus confirming Korean learners' preference for initial positions. All this leads to the conclusion that L1 transfer is at least partially (if not completely) responsible for the high frequency of misused sentence-initial *Especially* by L1 Korean writers. Despite some limitations and possible shortcomings, such as failure to examine L1 corpora of languages other than English and Korean, it is hoped that this study contributes to the development of learner corpus-based approaches to L2 acquisition, as well as the pedagogical application of corpora in the field of L2 writing instruction.

References

Anthony, L. 2024. *AntConc* (version 4.3.0) [Computer software]. Available online at https://www.laurenceanthony.net/software

Bell, D. M. 2007. Sentence-initial *and* and *but* in academic writing. *Pragmatics* 17(2), 183-201.

Bestgen, Y. 2014. Inadequacy of the chi-squared test to examine vocabulary differences between corpora. *Literary and Linguistic Computing* 29(2), 164-170.

Biber, D., S. Johansson, G. Leech, S. Conrad and E. Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Longman.

Boulton, A. 2010. Data-driven learning: Taking the computer out of the equation. *Language Learning* 60(3), 534-572.

Callies, M. 2015. Learner corpus methodology. In S. Granger, G. Gilquin and F. Meunier, eds., *The Cambridge Handbook of Learner Corpus Research*, 35-56. Cambridge: Cambridge University Press.

Carlstrom, B. and N. Price. 2012-2014. *The Gachon Learner Corpus*. Available online at http://koreanlearnercorpusblog.blogspot.kr/p/corpus.html

Davies, M. 2008-. *The Corpus of Contemporary American English (COCA)*. Available online at https://www.english-corpora.org/coca/

Davies, M. and D. Gardner. 2010. *A Frequency Dictionary of Contemporary American English: Word Sketches, Collocates and Thematic Lists*. London: Routledge.

Dunn O. J. 1964. Multiple comparisons using rank sums. *Technometrics* 6(3), 241-252.

Dupont, M. and S. Granger. 2022. *Connector placement in EFL learner writing: Focus on* however. Paper presented at the 6th Learner Corpus Research Conference.

Gablasova, D., V. Brezina and T. McEnery. 2017. Exploring learner language through corpora: Comparing and interpreting corpus frequency information. *Language Learning* 67(S1), 130-154.

Gilquin, G. and S. Granger. 2015. Learner language. In D. Biber and R. Reppen, eds., *The Cambridge Handbook of English Corpus Linguistics*, 418-436. Cambridge: Cambridge University Press.

Granger, S. 2012. How to use foreign and second language learner corpora. In A. Mackey and S. M. Gass, eds., *Research Methods in Second Language Acquisition: A Practical Guide*, 7-29. West Sussex: Wiley-Blackwell.

Granger, S. 2013. The passive in learner English: Corpus insights and implications for pedagogical grammar. *Learner Corpus Studies in Asia and the World* 1, 5-15.

Granger, S., E. Dagneaux, F. Meunier and M. Paquot. eds. 2009. *International Corpus of Learner English* (version 2). Louvain-la-Neuve: Presses Universitaires de Louvain.

Granger, S., M. Dupont, F. Meunier, H. Naets and M. Paquot. 2020. *The International Corpus of Learner English* (version 3). Louvain-la-Neuve: Presses Universitaires de Louvain.

Green, C. F., E. R. Christopher, J. Lam and K. Mei. 2000. The incidence and effects on coherence of marked themes in interlanguage texts: A corpus-based enquiry. *English for Specific Purposes* 19(2), 99-113.

Gries, S. T. 2015. Some current quantitative problems in corpus linguistics and a sketch of some solutions. *Language and Linguistics* 16(1), 93-117.

Huddleston, R. and G. K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Ishikawa, S. 2011. A new horizon in learner corpus studies: The aim of the ICNALE project. In G. Weir, S. Ishikawa and K. Poonpon, eds., *Corpora and Language Technologies in Teaching, Learning and Research*, 3-11. Glasgow: University of Strathclyde Press.

Ishikawa, S. 2023. *The ICNALE Guide: An Introduction to a Learner Corpus Study on Asian Learners' L2 English*. London: Routledge.

Johns, T. 1991. "Should you be persuaded": Two samples of data-driven learning materials. In T. Johns and P. King, eds., *Classroom Concordancing: ELR Journal* 4, 1-16. Birmingham: University of Birmingham.

Lee, S.-A. 2016. Verb agreement in English *wh*-clefts: A corpus-based study. *English Studies* 97(7), 779-806.

Lin, M. H. and J.-Y. Lee. 2015. Data-driven learning: Changing the teaching of grammar in EFL classes. *ELT*

Journal 69(3), 264-274.

Liu, Y. and S.-C. Rhee. 2017. On English adverb placements by Korean and Chinese EFL learners. *Studies in Linguistics* 45, 357-374.

McEnery, T. and R. Xiao. 2011. What corpora can offer in language teaching and learning. In E. Hinkel, ed., *Handbook of Research in Second Language Teaching and Learning*, 364-380. Abingdon: Routledge.

Nation, P. and D. Beglar. 2007. A vocabulary size test. *The Language Teacher* 31(7), 9-13.

National Institute of Korean Language. 2020. *NIKL Written Corpus* (version 1.2). Available online at https://kli.korean.go.kr/corpus

Nelson, G., S. Wallis and B. Aarts. 2002. *Exploring Natural Language: Working with the British Component of the International Corpus of English*. Amsterdam: John Benjamins Publishing Company.

Osborne, J. 2008. Adverb placement in post-intermediate learner English: A contrastive study of learner corpora. In G. Gilquin, S. Papp and M. B. Díez-Bedmar, eds., *Linking up Contrastive and Learner Corpus Research*, 127-146. Amsterdam: Brill.

Pyo, S.-Y. 2018. Functions of coordinators in native English and Korean EFL learner corpora. *English Language & Literature Teaching* 24(3), 45-67.

Quirk, R., S. Greenbaum, G. Leech and J. Svartvik. 1985. *A Comprehensive Grammar of the English Language*. New York: Longman.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing* (version 4.3.1) [Computer software]. Available online at http://www.R-project.org/

Rhee, S.-C. and C. K. Jung. 2014. Compilation of the Yonsei English learner corpus (YELC) 2011 and its use for understanding current usage of English by Korean pre-university students. *The Journal of the Korea Contents* Association 14(11), 1019-1029.

Richards, J. C. 1974. A non-contrastive approach to error analysis. In J. C. Richards, ed., *Error Analysis: Perspectives on Second Language Acquisition*, 172-188. Harlow: Longman.

Scheirer, C. J., W. S. Ray and N. Hare. 1976. The analysis of ranked data derived from completely randomized factorial designs. *Biometrics* 32(2), 429-434.

Sokal, R. R. and F. J. Rohlf. 1995. *Biometry: The Principles and Practice of Statistics in Biological Research* (3rd ed.). New York: Freeman.

Swan, M. 2005. *Practical English Usage* (3rd ed.). Oxford: Oxford University Press.

Tribble, C. 1989. The use of text structuring vocabulary in native and non-native speaker writing. *MUESLI News* Jun-89, 17-20.

Van Vuuren, S. and J. Berns. 2018. Same difference? L1 influence in the use of initial adverbials in English novice writing. *International Review of Applied Linguistics in Language Teaching* 56(4), 427-461.

Zhang, M. 2000. Cohesive features in the expository writing of undergraduates in two Chinese universities. *RELC Journal* 31(1), 61-95.

Examples in: English
Applicable Languages: English
Applicable Level: Tertiary

