



Definite and Indefinite Articles With Abstract Nouns in L2

Kent Lee

Pukyong National University

ARTICLE INFO

Received: 15 September 2020

Revised: 16 October 2020

Accepted: 15 November 2020

Examples in: English

Applicable Languages: English

Applicable Levels: Tertiary

KEYWORDS

definite articles/

indefinite articles/abstract nouns/

cognitive linguistics/

정관사/부정관사

추상명사/인지언어학

ABSTRACT

Lee, Kent. (2020). Definite and indefinite articles with abstract nouns in L2. *Modern English Education*, 21(4), 1-10.

Past studies have examined the problems that learners have with English definite, indefinite, and zero articles, particularly for learners whose L1 lacks an article system, such as Korean. Many of the empirical and theoretical studies of article acquisition have examined particular error types and types of nouns that are more problematic. In doing so, they have typically focused on more concrete nouns to the exclusion of abstract nouns. Abstract nouns tend to be used without articles, i.e., as bare nouns, a fact that learners may be aware of. However, such nouns can sometimes be used with articles for a more specific nuance. This study examines the difficulties that abstract nouns might pose for learners, as standard article rules may be unhelpful, especially for less familiar nouns. For this study, a linguistic survey with more abstract nouns was administered. The results showed that they had difficulties with more abstract noun patterns (as measured by semantic indices), and their difficulties were also affected by lexical frequency and complex noun phrases. Article patterns for more abstract nouns can be difficult to learn and teach, as their nuances can vary according to context. Some suggestions are sketched out for teaching such noun patterns.

I. INTRODUCTION

The English articles system has been a fairly productive area of L2 research, with learners whose L1 lacks article systems altogether (e.g., Korean, Chinese, and Russian), and languages with article systems that function differently enough to cause some errors in L2 English (e.g., Spanish, French, Greek, Vietnamese, or Hebrew). In recent years, more studies have been published that focus on East Asian learners. Many of these have been empirical studies using acceptability judgment tasks and similar methods. Some of these studies have examined semantic and pragmatic factors in the English article system, including those that pose challenges to Korean and other L2 learners, either as theoretical papers, or empirical papers investigating such linguistic factors.

However, these studies have focused primarily on non-abstract nouns, and often in isolated sentences with

no context beyond the sentence level. The earlier and more theoretical papers have mostly used concrete, physical nouns or animate nouns as examples of article usage patterns, and similarly, papers dealing with semantic and pragmatic factors or more theoretical aspects of L2 learning have tended to use physical or animate noun examples. This tendency has carried over into empirical studies of article acquisition, including those of East Asian learners. This can be seen in the types of questions and prompts used in acceptability judgment tasks, writing tasks, or elicitation tasks used in such studies. Few studies have examined more abstract nouns in detail, beyond theoretical or pragmatic studies of their use among native speakers. Researchers may shy away from more abstract nouns, as they are more difficult to analyze, test or research, while more concrete or animate nouns are more straightforward, and much easier to test or hypothesize about. However, such nouns are important in academic English, and are

Kent Lee (Professor)

Department of English Language and Literature, Pukyong National University, 45 Yongso-ro, Nam-gu, Busan, 48513, Korea

Tel: (051) 629-5385 / Email: kentlee7@gmail.com

important to students who need to learn academic writing or discourse styles for their studies or careers. It is thus important to better understand the types of errors and difficulties that students will likely encounter with more abstract noun phrases.

Abstractness is the converse of concreteness, where the latter refers to entities that can be directly perceived or experienced through one's physical senses, and thus, are physical, or tangible, form a physically distinct entity (entitativity), or can be mentally pictured (i.e., imageability). Abstractness is thus defined as referring to concepts that are not perceived through the senses, are less distinct, or are not readily imageable (e.g., Brysbaert, Warriner, & Kuperman, 2014). Nouns can thus be more or less concrete (e.g., *man*, *woman*, *TV*) or abstract (e.g., *peace*, *love*), along a continuum.

For this purpose, this study focuses on the more abstract types of nouns that one might encounter in more academic contexts. For such nouns, the traditional rules do not apply very well, specifically, the count/non-count rule (so-called countable nouns have articles, while mass nouns do not), and the first/subsequent mention rule (*a/an* for first mention, *the* for subsequent mentions). For example, one can speak of *literature* in a general sense, or *the literature of the British romantics* in contextually conditioned circumstances. In the next section, the relevant research is first surveyed, and then a linguistic survey study is described and the results are discussed below. The linguistic survey task was administered to Asian learners, focusing more on abstract nouns in humanities and social science contexts. As discussed below, abstract nouns are typically more difficult to analyze, but some research from psycholinguistics and cognitive linguistics can provide some useful insights for understanding their usage, learner errors, and teaching abstract noun patterns to learners.

II. LITERATURE REVIEW

Various theoretical discussions of English article rules have generally focused on data and examples of more concrete or physical nouns, namely, (1) animate and personal nouns for people, animals, or living agents; and (2) nouns for physical objects and materials. For example, earlier treatments in modern linguistics research by Hawkins focused on uniqueness constraints (Hawkins, 1980, 1991). An indefinite noun indicates a noun that is non-unique in context, e.g., "Fred lost a leg in the war" is acceptable, while "Fred lost *a nose in the war" (Hawkins, 1980, p. 56) is not; *a nose* is a unique entity while *a leg* is not. Definite nouns (with *the*) that have not been previously mentioned are possible with associative anaphors or associative definites, when it "instructs" the listener to identify a specific referent from a set of items. For example, "The only man at the party was Fred" (Hawkins, 1980, p. 55) is possible, as *party* instantiates the associated notion of attendees and the sentence identifies a particular attendee. In these papers, concepts such as uniqueness and associ-

ation are developed based on concrete nouns (e.g., *man*, *child*, *twin*, *professor*, *dog*, *paw*, *name*, *fact*, *party*, *leg*, *nose*). Lyons discusses similar principles of articles, whereby definites are licensed by discourse, situational, or set membership (Lyons, 1980), and invokes similarly concrete and animate nouns as examples (e.g., *door*, *light*, *students*, *house*, *book*, *newspaper*, *fact*, *person*, *maid*, *butler*).

This tendency has carried over into research on articles in second language acquisition. One well-known study (Liu & Gleason, 2002) examines learner difficulties (of various unspecified L1 backgrounds) with second-mention nouns, associative anaphors and associative nouns (e.g., *a wedding* and *the bride*), and overuse of *the* for general reference (*I think the tigers are beautiful animals*), grammatically null nouns (e.g., *chapter 20*), and so-called cultural or situational uses of bare nouns (*she had polio*). The test items in the appendix, however, are single-sentence items—isolated sentences with no context beyond the sentence level—which show a preponderance of physical and animate nouns (e.g., *window*, *computers*, *answer*, *handle*, *capital*, *pilot*) and a few nouns that are less concrete though not very abstract (e.g., *polio*). A study of children learning English as an L2 (Zdorenko & Paradis, 2008) showed that those from L1 backgrounds with no article systems (Chinese, Korean, and Japanese) had more difficulties and made more errors than children from languages with article systems (e.g., Spanish and Arabic), particularly in substituting *the* for *a* with first-mention nouns, or omitting articles. Children were asked to narrate stories from storybooks, which is fortunately more contextual, and unsurprisingly, the stimuli consisted of physical nouns (e.g., *ball*, *elephant*, *giraffe*, *lifeguard*). A verbal forced choice elicitation/discourse completion task with Chinese learners (Trenkic, 2008) showed that the participants tended to overuse *the* for indefinites and *a* for definites; the prompts seem to be mainly animate and object nouns (*girl*, *student*, *banker*, *president*, *owner*, *person*). Similarly conducted studies of learners of Serbian and other L1 backgrounds found that they tend to omit *the* from repeated occurrences of nouns, i.e., while they might use *the* for a second occurrence of a noun, they tend to omit it from subsequent mentions of the same noun, as they tend to find it redundant (Trenkic, 2002, 2009).

A few studies have included some more abstract nouns in their stimuli. One study of adult learners of Korean and other language backgrounds (Sun, 2016) showed that the indefinite *a* was learned first, then the definite *the* pattern, and zero articles were mastered last, as reflected in their error rates for these noun phrase types. This cloze completion study used mostly concrete nouns (e.g., *breakfast*, *omelet*, *phone*), but also some more fairly abstract nouns (e.g., *idea*, *presence*, *spirit*, *society*, *consciousness*), and this seems to be one study with a more diverse range of nouns. Another study of Japanese college students (Butler, 2002) used a cloze completion task followed by a retrospective interview and included concrete nouns as well as some conceptual nouns (e.g., *culture*, *idea*, *thought*). For example, one student explained his/her overuse of *the* (in "Japan

has *the old culture ...”) in that *culture* here referred to Japanese culture, which is specific, while another student omitted the indefinite article (“Japan has Ø culture ...”) and explained that *culture* is “not countable.” This shows an over-application of the non-count rule or a weak understanding of abstract nouns.

The general tendency for Korean learners to omit definite and indefinite articles is fairly well attested (Cowan, 2008; H. Lee, 1997), and similar patterns are attested for Chinese learners (Robertson, 2000). Koreans, for example, tend to treat some non-count as countable, and then using them as indefinite singulars or bare plurals (e.g., **equipments*, **an equipment*, **researches*, **informations*), possibly due to transfer effects (Cowan, 2008). Learners are also likely to have difficulty with cataphoric uses of *the* in a new, first-mention noun, and situational uses (I. W. Yoo, 2009). In cataphora, the sentence itself introduces and identifies the referent, especially when the definite noun is followed by a post-modifier that identifies it, as in (1a). Situational uses are especially common in conversation, where speakers treat other objects in the scene as definite, as in (1b).

(1a) The book that she bought was stolen.

(1b) As she entered, the phone was ringing and the TV was on.

As I. W. Yoo (2009) explains, these uses are generally not treated in the ESL/EFL grammar books or textbooks that are produced in Korea or that are aimed at Koreans. However, a few general ESL/EFL texts for general audiences do discuss definite nouns that are made definite by post-modifier expressions as in (1a); this includes a few applied linguistics texts for learners (Swan, 2011), and a few texts on academic and technical writing (Huckin & Olsen, 1990). Such post-modifiers often consist of relative clauses (as above) and prepositional phrases (e.g., *the theory of relativity*). Also, Korean writers tend to drop articles when a noun phrase contains two or more adjectives between the article position and the noun (e.g., *Ø *heavy red book*), as opposed to simple, unmodified noun phrases (*the book*) (H. Lee, 1999).

Studies of Korean and Chinese learners show some interesting patterns with the aforementioned associative nouns, i.e., first-mention definite nouns that are related to previously mentioned entities. Such nouns are those that native speakers would mark these as definite, including category based anaphora (e.g., *the novel* → *the book*), and bridging inferences based on world knowledge (e.g., *baked* → *the cake* or *birthday* → *the cake*). Learners tend not to treat these as definite on first mention, as shown with proficiency and acceptability judgment tasks with Korean students (J. Cho, 2017), and studies of writing samples, acceptability judgment tasks, and corpora of Chinese learners (Chan, in press; Feng, 2019; Robertson, 2000). In these studies, the nouns or stimuli seem to be mostly physical nouns (e.g., *cake*, *dessert*, *microwave*). Koreans also tend to have difficulty interpreting bare plurals as gener-

ics, and some participants instead interpret bare plurals as definites (e.g., *Cats are independent*) or definite plurals (*the books*) as generic (Ionin & Montrul, 2009). Another study (H. Ko, Ionin, & Wexler, 2010) found that Koreans tend to overuse *the* in contexts where the existence of presuppositional of an indefinite reference cannot be assumed; for example, in a sentence like *After the game I went to talk to a player/*the player/one of the players*, the phrase *a player* is acceptable since a particular player cannot be presupposed, except in a partitive expression (one of the players) that identifies a player; yet some Korean learners tend to use *the player* in such contexts, more like an associative definite noun.

Several studies by Ionin and colleagues have found that the semantic property of specificity can be problematic, that is, whether a noun is specifically known to at least the speaker regardless of its status as a first-mention indefinite or subsequently mentioned definite. For example, in the sentence *My room has a desk*, the noun *desk* is specific (not just any desk) and known to the speaker, but indefinite as a first-mention noun. One study (Ionin, H. Ko, & Wexler, 2004) of Korean and Russian learners performing forced elicitation and production tasks found that learners tend to overuse *the* for first-mention specific nouns. Similar results were reported for Korean and Russian learners with reading and semantic judgment tasks, in that they tended to overuse *the* for nouns that were specific but non-definite (e.g., *She wants to buy a book* in a context where she wants a particular book) (L. K. Kim & Lakshmanan, 2009). Another study with elicitation and comprehension tasks also shows that Koreans are not always able to distinguish definite and specific nouns (Ionin, S. Baek, E. Kim, H. Ko, & Wexler, 2011). Similar patterns of overuse of *the* with specific indefinites have been reported for Japanese and Russian learners (Trenkic, 2009, and references therein).

From the previous research, we know that learners from Korean and other L1 backgrounds have difficulty with general omission of articles, generic noun phrases, associative noun phrases, at least some complex noun phrases, and cataphoric noun phrases, such as those followed by a post-modifier that identifies the noun. In the previous studies, the stimuli were largely physical or animate nouns, and few abstract nouns were involved in the various tasks given to participants. We might expect similar patterns for abstract nouns, but we might also expect greater difficulty for abstract nouns that are post-modified and identified as definite (e.g., *the theory of relativity*). Also, as best as can be discerned, the test items in most of the aforementioned studies often consist of individual sentences with little context. Since article use depends crucially on context, longer contexts would be preferable, for example, in an elicitation or acceptability survey consisting of longer reading passages. This leads to the research questions of this study:

- 1) Do similar error patterns emerge for abstract nouns? For example, do they have similar problems with first-mention nouns that are made specific by the

context, associative nouns, or more complex noun phrases?

- 2) Since more abstract nouns can be used with more abstract or more specific senses, particularly post-modified nouns (e.g., *theory*, *a theory*, *the theory of relativity*), do they have more difficulty with such nouns when used with different meanings?

To examine these patterns, a contextualized fill-in-the-blank survey study is reported in the next section.

III. METHOD

1. Instrument

A contextualized forced-choice survey was designed with 110 test items, divided into three passages consisting of three or more paragraphs each for a meaningful context. The passages were written by the researcher, and dealt with (1) literary theory (with 47 noun phrases as test items), (2) paralinguistics and linguistics (45 noun phrases), and (3) computing (18 noun phrases), as these topics seemed relevant to the target audience, and allowed for a large number of abstract nouns. Blanks were inserted before each of the target nouns, and participants were instructed to fill in the blank with an appropriate form: (1) the indefinite *a/an*; the definite *the*; or the null symbol $\emptyset$ for no article.

2. Participants and Procedure

The participants were 20 undergraduate students in an upper-level undergraduate English linguistics course, who filled out the printed survey for course credit as a minor assignment (and provided informed consent as participants). Basic demographic information was collected, such as gender, major, language background, and amount of time spent living abroad in Anglophone countries. Of the participants, sixteen were majoring in English or related fields; three were in business or economics; and one in media studies. Thirteen were female and seven were male; fourteen were Korean and four were Chinese, plus one Russian and one Vietnamese. Fourteen had not spent a significant amount of time in Anglophone countries; four had spend 0.5 to 1 year abroad; and two had spent 3-6 years abroad. Their responses were compared to the correct or pragmatically felicitous answers, as determined by the researcher (a native English speaker), and identified as correct or incorrect.

3. Analysis

Syntactic and semantic factors were included as independent variables. Syntactic properties of the noun phrases were categorized and entered as variables: (1) whether the

noun phrase is a simple, unmodified noun, or is pre-modified by one or more adjectives before the noun (e.g., *a heavy red book*); (2) whether it is followed by a post-modifier that modifies or identifies the noun (namely, a prepositional phrase, relative clause, modifying infinitive phrase, or participial phrase); (3) NP or nominal phrase type (simple, compound, or coordinate noun phrase, such as *volume and intonation*), but this factor was not significant in the statistical analysis; and (4) noun number (singular or plural). These were entered as categorical variables. To better examine the semantic effects of abstractness versus concreteness, semantic indices were included as independent variables. A set of semantic ratings for 40,000 words created by psycholinguistics researchers (Brysbaert et al., 2014) was used, consisting of large-sample normed ratings from native English speakers. The normed data consist of ratings from 1.0 (abstract) to 5.0 (concrete). The Brysbaert data set also includes lexical frequencies from large corpus data, with a lexical frequency for each word expressed as a natural logarithmic score. One contextual variable was included: how many times the noun had occurred thus far in the text (sequential occurrence or xth occurrence), e.g., first mention, second mention, etc. Since lexical data are involved, which do not follow a standard distribution, logistic regression models were used to identify relevant factors. Such tests include logarithmic based distributions (Agresti, 2002), which are better suited for lexical data. Parametric tests like ANOVAs and *t*-tests, even with adjustments, can yield false negatives or less reliable results with data involving lexical frequency data and other nominal count data (Baayen & Lieber, 1996). Logistic and logarithmic regression are also more robust to unbalanced counts, as in linguistic data where some data cells or variable combinations may be much larger than others. They are thus ideal for categorical lexical data.

IV. RESULTS

With many abstract nouns, nouns repeated with sometimes varying nuances, and fairly academic texts, the task seemed somewhat difficult, as the participants exhibited an overall error rate of 37.8% (832 errors out of 2200 responses). The descriptive data are shown in Table 1, for the expected or correct response (*a/an*, *the*, $\emptyset$). Some respondents occasionally indicated two possible responses (*the/a* or *the/∅*). The error rate was highest for nouns that should be marked as definites, followed by nouns that should be bare noun phrases (zero article nouns) and indefinite noun phrases.

The various factors were entered into a logistic regression model, since the dependent variable is binary (i.e., whether the response was correct or incorrect); logistic regression was carried out in Proc Catmod in SAS 9.4 for binary response variables. The statistically significant factors were as follows: L1 background; pre-modification (simple nouns versus modification by one or more adjectives); post-modification (identifying relative claus-

es, prepositional phrases, infinitive phrases, or participle phrases), sequential occurrence, and lexical frequency. Surprisingly, semantic abstractness/concreteness was not significant, nor was NP type (simple, coordinate, compound). The output of the logistic regression shows estimates for relevant levels of categorical variables, as well as odds ratios. These are shown in Table 2, where the effect sizes and odds ratios model the likelihood of errors; categorical variable levels that were significant or approached significance are shown.

TABLE 1
Correct and Incorrect Responses

NP type (expected response)	n	Correct	Incorrect	Error rate	Total (%)
Ø (zero article)	68	857	503	37.0	1360 (61.8)
the	26	305	216	41.5	521 (23.7)
a/an	14	170	110	39.3	280 (12.7)
the/a	1	16	3	15.8	19 (0.9)
the/Ø	1	20	0	0.0	20 (0.0)
Total	110	1368	832	37.8	2200 (100.0)

The L1 background seemingly had some effect, mainly in that the Korean students were slightly less likely to make errors than the Chinese students, but more likely than the other L1s (one Russian, one Vietnamese¹), but a comparison with these other L1s is unreliable or not possible due to the very small sample size. The Koreans' slight advantage may be due to more exposure to English, be it from time abroad or more years of English in school. However, since time abroad as a factor was not statistically significant here, perhaps because those who performed better had more exposure to English from international schools in their home countries or from other sources. One or both of these had a significant effect here, but since the role of L1 and exposure are unclear here, such factors warrant further future research to disambiguate these effects.

Simple noun phrases seemed generally easier for participants, as participants were 1.4 times more likely to make

errors with nouns modified by one adjective, according to the odds ratio, and 1.8 times as likely to make errors on nouns with two preceding adjectives. Post-modification was more problematic, as they were only 0.8 times as likely to make errors with simple noun phrases than with nouns followed by a modifier. Sequential occurrence had a significant effect. Some nouns (e.g., *criticism* in the abstract sense of literary criticism, and *theory*) were repeated often in the text, and other nouns a few times. Respondents showed a marginally significant tendency (1.3 times more likely) to make errors on the first mention, then less likely so on fourth and sixth mentions, but much more likely to make errors ($\theta = 12.3$) on the seventh mention. As seen below, errors came in the form of inserting unnecessary articles on abstract nouns. Other positions (2nd mention, etc.) were not significant, probably due to a limited sample size with only a few oft-repeated nouns. It seems that multiple mentions were sometimes helpful in making correct responses, but up to a point. If many mentions do indeed tend to induce more errors (e.g., by the twelfth mention), this could be an effect of semantic satiation, or the high discourse prominence of the noun, such that the respondents might have viewed *the* as redundant, as in the Trenkic (2002, 2009) studies of Serbian and other learners. Singular nouns seemed easier, as they only 0.87 times as likely to make mistakes on singular noun phrases than plurals. Lexical frequency was advantageous, and conversely, low frequency nouns were more difficult and error-prone. More frequent nouns were more familiar to respondents, and they were only 0.86 times as likely to make mistakes with more frequent nouns than with less frequent nouns. However, since the lexical frequencies from Brysbaert et al. (2014) are natural logarithms, more research is needed to better quantify these frequency effects in a manner that would be meaningful, say, for language educators.

Some errors are shown in the following item samples. The item number appears in superscript before the underlined correct response (*the*, *a/an*, *Ø*) and before the

TABLE 2
Analysis of Response Accuracy

Factor	Level	Effect (β)	Odds ratio (θ)	df	Chi-square	Probability
L1	(Kor. cf. Ch.)	-1.1572	0.314	3	64.57	< .0001
Pre-modification	overall			4	20.06	0.0005
	1 Adj.	0.3270	1.387		3.87	0.0491
	2 Adj.	0.5606	1.752		6.29	0.0121
Post-modification	(unmodified)	-0.1675	0.846	1	10.37	0.0013
Sequential occurrence	overall			8	51.05	< .0001
	1x	0.2798	1.323		3.29	0.0698
	4x	-0.5856	0.557		5.65	0.0174
	6x	-0.5719	0.564		3.49	0.0619
	7x	2.5071	12.269		14.58	0.0001
Number (sg/pl)	(singular)	-0.1454	0.865	1	4.68	0.0305
Lexical frequency	(higher freq.)	-0.1527	0.858		17.62	< .0001

¹ The Vietnamese student had higher response accuracy rates, probably due to international experience, and the fact that Vietnamese has some type of article system, but no specific comparison is attempted here due to the small sample size.

corresponding underlined noun, and in square brackets counts are given for erroneous responses.² For example, for item 101, the correct answer is *Ø* (i.e., *Ø Literary criticism*), with one participant's incorrect insertion of *the*, and for item 102, the correct answer is *Ø literature*, with five incorrect participant responses with *the*. For the abstract nouns *criticism*, *literature*, *theory*, we see a tendency to mark them with definite or indefinite articles at times throughout these examples, even though these nouns are continually used in an abstract sense. Thus, they were sometimes confused about the intended nuance, or the optimal form to use. We also see these terms marked with *a* or *the* by respondents when modified by one or more adjectives. In item 111 we have an example of an associative definite, which refers to the preceding terms of the text, yet nine respondents preferred the zero article. In item 115 we see *works* used as a plural indefinite generic, yet six respondents marked it with *the* and three with *a*, indicating possible confusion in their interpretation; perhaps they assumed the noun referred to specific works, though that is not supported by the context. In item 124, we see the noun *century* made definite by *4th*, yet two respondents preferred a zero article.

(2a) Section 1 examples

¹⁰¹*Ø* [*the: 1] Literary ¹⁰¹*criticism* is the study, evaluation, and interpretation of ¹⁰²*Ø* [*the: 5] ¹⁰²*literature*. ¹⁰³*Ø* [*the: 3; *a: 5] Modern literary ¹⁰³*criticism* is often informed by ¹⁰⁴*a* [*the: 2; *a: 3] literary ¹⁰⁴*theory*, which is ¹⁰⁵*a* [*Ø: 2; ?the: 8] philosophical ¹⁰⁵*discussion* of its methods and goals. Though the two activities are closely related, literary critics are not always, and have not always been, theorists. Whether ¹⁰⁶*Ø* [*the: 6; *a: 1] literary ¹⁰⁶*criticism* ought to be considered a separate field of inquiry from ¹⁰⁷*Ø* [*the: 7; *a: 6] literary ¹⁰⁷*theory*, or conversely from book reviewing, is ¹⁰⁸*a* [*the: 3] ¹⁰⁸*matter* of some controversy. For example, X draws no distinction between ¹⁰⁹*Ø* [*the: 7; *a: 5] literary ¹⁰⁹*theory* and ¹¹⁰*Ø* [*the: 5; *a: 2] literary ¹¹⁰*criticism*, and almost always uses ¹¹¹*the* [*Ø: 9] ¹¹¹*terms* together to describe the same concept. Some critics consider ¹¹²*Ø* [*the: 5; *a: 2] literary ¹¹²*criticism* a practical application of ¹¹³*Ø* [*the: 6; *a: 2] literary ¹¹³*theory*, because ¹¹⁴*Ø* [*the: 7; *a: 3] ¹¹⁴*criticism* always deals directly with ¹¹⁵*Ø* [*the: 6; *a: 3] particular literary ¹¹⁵*works*, while ¹¹⁶*Ø* [*the: 9; *a: 4] ¹¹⁶*theory* may be more general or abstract.

(2b) Section 1 examples

In ¹²⁴*the* [*Ø: 2] ¹²⁴*4th* ¹²⁴*century* BC, Aristotle wrote the *Poetics*, a typology and description of ¹²⁵*Ø* [*the:

7] literary ¹²⁵*forms* using many specific criticisms of ¹²⁶*Ø* [*the: 6] contemporary ¹²⁶*works* of ¹²⁷*Ø* [*the: 4; *an: 2] ¹²⁷*art*. Later classical and medieval criticism often focused on ¹²⁸*Ø* [*the: 8] religious ¹²⁸*texts*, and ¹²⁹*Ø* several long religious ¹²⁹*traditions* of hermeneutics and textual exegesis have had ¹³⁰*a* [*Ø: 5; *the: 1] profound ¹³⁰*influence* on the study of ¹³¹*Ø* [*the: 8] secular ¹³¹*texts*. This was particularly ¹³²*the* [*Ø: 2; *a: 5] ¹³²*case* for ¹³³*the* [*Ø: 12] literary ¹³³*traditions* of the ¹³⁴*Ø* [*Ø: 9;] three Abrahamic ¹³⁴*religions*: ¹³⁵*Ø* [*the: 8; *a: 1] Jewish ¹³⁵*literature*, ¹³⁶*Ø* [*the: 7; *a: 1] Christian ¹³⁶*literature* and ¹³⁷*Ø* [*the: 8; *an: 1] Islamic ¹³⁷*literature*.

(2c) Section 2 examples

²¹⁴*A* [*the: 14] ²¹⁴*film* that Gumperz made for ²¹⁵*the* [*Ø: 18] ²¹⁵*BBC* in 1982, does ²¹⁶*a* [*Ø: 1] particularly good ²¹⁶*job* of demonstrating ²¹⁷*the/Ø* cultural ²¹⁷*differences* in paralanguage, and ²¹⁸*the* [*Ø: 2; *an: 1] ²¹⁸*impact* these have on relationships.

Finally, we see examples of post-modified nouns causing confusion in Table 3, as they are variously intended as definite, indefinite, or abstract generic nouns, depending on the intended nuance in the context. In the following examples, some post-modified nouns are indefinite (e.g., *a good job of demonstrating*), where the post-modifier simply serves as a descriptor or as an explanatory phrase, rather than as an identifier; some are definite (e.g., *the impact these have*); and some are plural generics (e.g., *works of art*).

TABLE 3
Examples of Errors With Post-Modified Noun Phrases

Item	Text	Errors	Modifier
108	<i>a matter</i> of some controversy	*the: 3	Prep
126	<i>Ø</i> contemporary <i>works</i> of art	*the: 6	Prep
129	<i>Ø</i> religious <i>traditions</i> of hermeneutics	*the: 4	Prep
130	<i>a</i> profound <i>influence</i> on the study	*Ø: 5; *the: 1	Prep
132	<i>the case</i> for the literary traditions	*Ø: 2; *a: 5	Prep
133	<i>the</i> literary <i>traditions</i> of the three religions	*Ø: 12	Prep
214	<i>A film</i> that Gumperz made	?the: 14	Relative clause
216	<i>a</i> particularly good <i>job</i> of demonstrating	*Ø: 1	Prep
218	<i>the impact</i> these have on relationships	*Ø: 2; *an: 1	Reduced relative clause

² Naturally, asterisks indicate erroneous or infelicitous responses, and question marks indicate responses that would be dispreferred or less fitting in the context. Some noun phrases were not evaluated, such as proper nouns (*the Poetics*), serial compound noun phrases (*the study, evaluation and interpretation of ...*), and cardinal numerals (*the two activities*), since these were not of interest, were too infrequent to compare, or were lacking in the semantic indices used here. Sentence-initial nouns preceded by blanks were capitalized, regardless of the correct or expected answers, since sentence beginnings were viewed as relevant contextual cues that were not of interest here.

Of the 43 post-modified nouns queried, seven were indefinite singular noun phrases in context, 21 were definite (four singular, 17 plural), seven were singular bare nouns, seven were plural bare nouns, and one could be either definite or bare plural. The error rates for these are shown in Table 4, where we see that post-modified noun phrases intended as plural definites (e.g., *the traditions of these religions*) in context are most problematic, though these were few in number, followed by bare nouns that are intended as singular abstract nouns (e.g., *literary criticism of the Renaissance*) or plural generics (e.g., *works of art*, where the head noun *works* is a generic plural and the entire phrase functions semantically as a whole).

TABLE 4
Error Rates for Post-Modified Noun Phrases

Expected form	Number	Correct	Incorrect	Total responses
a	singular	110 (78.6%)	30 (21.4%)	140
the	singular	214 (62.9%)	126 (37.1%)	340
	plural	29 (36.3%)	51 (63.8%)	80
the/Ø	plural	20 (100.0%)	0 (0.0%)	20
Ø (bare noun)	singular	74 (52.9%)	66 (47.1%)	140
	plural	96 (68.6%)	44 (31.4%)	140
Total		543 (63.1%)	317 (36.9%)	860

The results show some evidence for difficulty with abstract noun phrases, particularly post-modified NPs, whose interpretation can vary according to context and intended usage. But since abstractness/concreteness was not significant, the effects of abstractness seem unclear here. This may be because the survey had a preponderance of abstract nouns, making it difficult to statistically assess its influence. Some factors warrant further research, such as L1 background, L2 exposure, L2 proficiency levels, and other demographic data that were not available here. In both models, NP phrase type (simple, coordinate, compound NP) was not significant, though this may be because the survey had few tokens of coordinate and compound NPs. Some noun phrase types could not be examined, such as those with various quantifiers. Types of post-modification could not be compared here (e.g., differential effects of relative clauses versus prepositional phrases, among others). Thus, future research with more participants, demographic information, a larger survey and a greater variety of concrete to abstract nouns, and a greater variety of noun phrase types, is needed.

V. DISCUSSION

The results of the survey task showed that abstract nouns may be problematic, even when nouns are intended fairly unambiguously as abstract nouns in texts. The Korean and other learners in this task tended to mark them as definite or indefinite nouns instead, especially when the nouns were used repeatedly but still in an abstract sense. It is possible that they may be confused by textbook first/

subsequent mention rule that instructs learners to mark subsequent mentions of a noun with *the*, versus their likely intuitions that the noun is abstract. The pattern here for repeated abstract nouns is reminiscent of, but possibly more complex than the reported tendency for some learners to omit articles on definite nouns that are repeated (Trenkic, 2009). Here the nouns were more abstract, where some of them would be continuously used as bare nouns (e.g., multiple instances of *Ø criticism*), and a greater variety of errors was seen here with repeated nouns, such as article insertions on bare nouns (**a/*the criticism*), and article deletions on definite and indefinite nouns (*the impact* → **Ø impact*).

Nouns, including abstract nouns, can also be more difficult for learners when the noun phrases are more complex, e.g., when the noun is modified by preceding adjectives, or especially when the noun is post-modified. The likelihood of errors with nouns preceded by one or more adjectives here is consistent with previous studies showing a similar tendency for Korean learners (H. Lee, 1999). Post-modified nouns are understandably confusing, since the syntactic structure itself does not provide salient cues about whether the noun should be interpreted as abstract, generic, definite, or indefinite. An abstract noun (e.g., *theory* or *art*) can be post-modified to indicate a less abstract, more specific, and more definite meaning (e.g., *the theory of relativity*). But post-modification itself is not necessarily a reliable cue, as a post-modified noun can be abstract (e.g., *we are studying Ø contemporary works of art*); it can be indefinite, for example, if the entire complex noun phrase is descriptive or definitional (e.g., *it's a type of therapy*); or it can easily be made definite on first mention when the modifier serves to identify a particular type or entity (*the theory of relativity* or *the art of Van Gogh*).

Lexical frequency had a facilitatory effect, in that more frequent words tended to be less error prone, and less familiar words were more error-prone (e.g., *intonation*, *prosody*, *utterance*), as the learners were less familiar with their usage patterns. Plural nouns were slightly more difficult, particularly post-modified plurals (e.g., *works of art*) as it may be difficult to discern whether the noun phrase should be interpreted as abstract, definite, or indefinite (e.g., *the stars* = the stars in the sky). The survey materials in this study had relatively few associative nouns, and this deserves more detailed study. A few associative nouns did exhibit some errors, e.g., when *the terms* (item 111) referred to the terminology of the preceding context, or when *the case* (item 132) referred to the previously described situation, some respondents gave the zero article (Ø) as a response. It is likely that learners from L1s without articles would have difficulty with associative nouns that are less physical or concrete, so further study is needed.

Interestingly, lexical semantic concreteness did not show up as a significant factor in the statistical model. This may be due to the limited size of the study (only 20 participants), or more likely, due to a lack of variety in the test stimuli, as the survey was designed with a heavy pre-

ponderance of abstract nouns. Further research is needed with a larger, more diverse set of survey items and participants. Since the abstract nouns in this study were difficult for the learners, with a 37% error rate, a follow-up study with a more diverse range and number of nouns ranging from concrete to abstract would likely show more overall difficulties with abstract nouns.

In contrast to past studies, this study shows difficulties for some types of abstract nouns, namely, post-modified nouns, and repeated mentions of nouns. Since the survey items are more contextualized, this seems to have more validity, since article usage depends on context and speaker intention. As mentioned, this study has some limitations. Some demographic information such as L2 proficiency level was lacking, and a greater variety of noun phrases, concrete and abstract nouns, and survey items would be helpful here. The statistical model allowed for examination of noun phrase types, but a larger data set would help to examine the effects of semantic factors, proficiency levels and L1 background. These will be included in future follow-up research on abstract nouns.

The pedagogical implications of this study, albeit fairly exploratory, seem clear. As H. Lee (1999) and I. W. Yoo (2009) have pointed out, the traditional ESL/EFL textbooks, materials, and textbook rules are inadequate and offer insufficient guidance, particularly the count/non-count noun rule and the first/subsequent mention rule. This can be particularly problematic with abstract nouns that do not follow these rules very clearly. Abstract nouns can easily exist as zero article bare nouns throughout multiple mentions, or can be made more specific by post-modifiers (e.g., *a theory of gravity* or *the theory of gravity*). More practice with realistic or authentic texts and with contextualized examples is needed. Exercises need to be contextualized in paragraph or essay-level contexts, not the isolated, individual, context-free sentences that are the staple of many grammar books, since articles depend crucially upon context. Students need to engage with more natural texts or discourse samples, in which nouns are used differently, especially less concrete and more abstract nouns. Since helpful guidance in the form of clear explanations of abstract noun phrases, it will be helpful if students can engage with texts under the guidance of teachers who can explain the context and intended nuances of nouns. That is, authentic texts should be a priority for more advanced students. At higher levels, the traditional rules of articles may need to be abandoned, and instead students need practice in reading and writing texts, and in discerning the contextual and pragmatic cues that might render a noun generic, abstract, indefinite, or definite. This might give students a better understanding that such noun phrases might refer to a fully abstract concept (e.g., *feminism*), a generic concept (e.g., *works of art*), or a specific instance or type of the concept (e.g., *the feminism of the 1960s* or *the art of the Renaissance*).

Both implicit and explicit instruction may be helpful for intermediate and advanced learners, such as the aforementioned guided discussion of authentic texts. Integrated

reading and writing instruction based on such texts may be helpful, as it has been found to be helpful for general grammatical and vocabulary learning (e.g., J. K. Kim & E. J. Moon, 2019). However, this researcher's contention is that articles are not strictly grammatical, but depend on context and speaker intention, especially with more abstract nouns. Thus, it may be helpful to treat article problems more as pragmatic errors than grammatical, since pragmatic errors may be more amenable to correction than grammatical errors at the college level, as W. Jung's (2019) study would suggest. Since learners may dislike grammar, de-emphasizing grammar and emphasizing pragmatic context may be more successful, and such an approach deserves further research with article acquisition. Thus, guided discussion can help students to guess writers' expressive intentions in their selection of articles in noun phrases, and to be more aware of their intentions when they express themselves in writing, e.g., by responding to a piece of writing they have discussed.

Unfortunately, practical advice for learners is lacking regarding these more complex situations with associative, cataphoric, and post-modified nouns. Much more research is needed in order to uncover helpful cues or patterns that can be taught to learners. In the absence of such guidance, learners require much more guided or individual practice with more naturalistic, academic, or complex texts and discourse, so that they can develop a better sense of the contextual cues and pragmatic cues that affect the preferred type of noun phrase in such contexts.

VI. CONCLUSION

This study has shown that learners from L1s lacking article systems have considerable difficulties with less concrete and more abstract nouns. They particularly have difficulty with more complex noun phrases, such as post-modified nouns, and the different ways that a noun can be used in context.

This initial study requires much more detailed follow-up study, as some semantic and pragmatic patterns could not be examined in sufficient detail here. A larger sample size with a wider variety of noun phrases is needed, so that different error types can be studied and compared, and so that more factors and interaction effects can be examined (e.g., singular/plural, noun phrase types, associative nouns, cataphoric nouns, semantic variables). More comparisons and analysis are needed with associative nouns, post-modified nouns, nouns with a wider range of differing lexical frequencies, and nouns with a wider range of abstractness and concreteness, and thus, a wider range of nouns needs to be tested. Also, research is specifically needed on nouns that might shift in nuance and article usage within one text, e.g., a text in which abstract nouns like *theory* are used interchangeably (*gravitational theory* cf. *a theory of gravity* cf. *the theory of gravity*). Similarly, the effects of polysemy (multiple possible meanings of nouns) and article patterns require investigation, as this

has received little attention in the research literature. For example, *text* can refer to a text message or a book (*a text*), or slightly more abstractly to a collection of words in print (*the advantages of text*). This initial study of abstract nouns will hopefully lead to a more thorough investigation with larger data sets, so that we can provide learners and teachers with tools for a better understanding of English noun phrase patterns.

REFERENCES

- Agresti, A. (2002). *Categorical data analysis*. New York: John Wiley & Sons.
- Baayen, R. H., & Lieber, R. (1996). Word frequency distributions and lexical semantics. *Computers and the Humanities*, 30(4), 281-291.
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46(3), 904-911.
- Butler, Y. G. (2002). Second language learners' theories on the use of English articles: An Analysis of the metalinguistic knowledge used by Japanese students in acquiring the English article system. *Studies in Second Language Acquisition*, 24(3), 451-480.
- Chan, A. Y. W. (in press). Typology and contexts of article errors: Investigation into the use of English articles by Hong Kong Cantonese ESL learners. *International Review of Applied Linguistics in Language Teaching*. Retrieved from <https://doi.org/10.1515/iral-2018-0268>
- Cho, Jacee. (2017). The acquisition of different types of definite noun phrases in L2-English. *International Journal of Bilingualism*, 21(3), 367-382.
- Cowan, R. (2008). *The teacher's grammar of English*. New York: Cambridge University Press.
- Feng, S. (2019). The acquisition of English definite noun phrases by Mandarin Chinese speakers. *Studies in Second Language Acquisition*, 41(4), 881-896.
- Hawkins, J. A. (1980). On surface definite articles in English. In J. van der Auwera (Ed.), *The semantics of determiners* (pp. 41-66). Baltimore, MD: University Park Press.
- Hawkins, J. A. (1991). On (in)definite articles: Implicatures and (un)grammaticality prediction. *Journal of Linguistics*, 27(2), 405-442.
- Huckin, T. N., & Olsen, L. A. (1990). *Technical writing and professional communication for nonnative speakers of English*. New York: McGraw-Hill.
- Ionin, T., Baek, Soondo, Kim, Eunah, Ko, Heejung, & Wexler, K. (2011). That's the meaning: Interpretation of definite and demonstrative descriptions in L2-English. In M. Pirvulescu, M. C. Cuervo, A. T. Pérez-Leroux, J. Steele, & N. Strik (Eds.), *Selected Proceedings of the 4th Conference on Generative Approaches to Language Acquisition North America (GALANA 2010)* (pp. 122-138). Somerville, MA: Cascadia Press.
- Ionin, T., Ko, Heejeong, & Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition*, 12(1), 3-69.
- Ionin, T., & Montrul, S. (2009). Article use and generic reference: Parallels between L1-and L2-acquisition. In M. D. P. G. Mayo & R. Hawkins (Eds.), *Second language acquisition of articles: Empirical findings and theoretical implications* (pp. 147-174). Philadelphia: John Benjamins.
- Jung, Woo-hyun. (2019). L2 learners' pragmatic awareness in comparison to grammatical awareness in tertiary education settings. *Modern English Education*, 20(2), 19-32.
- Kim, Jae Kyung, & Moon, Eun Joo. (2019). The effect of integrating reading and writing on EFL learners' writing ability and vocabulary usage. *Modern English Education*, 20(3), 57-68.
- Kim, Lucy Kyoungsook, & Lakshmanan, U. (2009). The processing role of the Article Choice Parameter. In M. D. P. G. Mayo & R. Hawkins (Eds.), *Second language acquisition of articles: Empirical findings and theoretical implications* (pp. 87-114). Philadelphia: John Benjamins.
- Ko, Heejeong, Ionin, T., & Wexler, K. (2010). The role of presuppositionality in the second language acquisition of English articles. *Linguistic Inquiry*, 41(2), 213-254.
- Lee, Hikyoung (1997). English article deletion in Korean EFL learners' compositions. *Working Papers in Educational Linguistics*, 13(2), 41-50.
- Lee, Hikyoung (1999). Variable article use in Korean learners of English. *University of Pennsylvania Working Papers in Linguistics*, 6(2), 35-47.
- Liu, D., & Gleason, J. L. (2002). Acquisition of the article by nonnative speakers of English: An analysis of four nongeneric uses. *Studies in Second Language Acquisition*, 24(1), 1-26.
- Lyons, C. G. (1980). The meaning of the English definite article. In J. van der Auwera (Ed.), *The semantics of determiners* (pp. 81-95). Baltimore, MD: University Park Press.
- Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research*, 16(2), 135-172.
- Sun, G. (2016). The acquisition of English articles by second language learners: The sequence, differences, and difficulties. *SAGE Open*, 6(1), 1-8.
- Swan, M. (2011). The definite article. *New Scientist*, 212(2834), 33.
- Trenkic, D. (2002). Establishing the definiteness status of referents in dialogue (in languages with and without articles). *University of Cambridge Working Papers in English and Applied Linguistics*, 7(1), 107-131.
- Trenkic, D. (2008). The representation of English articles in second language grammars: Determiners or adjectives? *Bilingualism*, 11(1), 1-18.
- Trenkic, D. (2009). Accounting for patterns of article omissions and substitutions in second language production. In M. D. P. G. Mayo & R. Hawkins (Eds.), *Second language acquisition of articles: Empirical*

findings and theoretical implications (pp. 115-143).
Philadelphia: John Benjamins.

Yoo, Isaiah WonHo. (2009). The English definite article: What ESL/EFL grammars say and what corpus findings show. *Journal of English for Academic Purposes*, 8(4), 267-278.

Zdorenko, T., & Paradis, J. (2008). The acquisition of articles in child second language English: Fluctuation, transfer or both? *Second Language Research*, 24(2), 227-250.