

# L2 Article Semantics and Second Language Processing

Kyoungsook Kim and Usha Lakshmanan

University of Southern California and Southern Illinois University at Carbondale

## 1. Introduction

The English article system is generally known to pose learnability difficulties for adult learners of English as a Second Language (L2). Even to advanced L2 learners who demonstrate a strong command in both spoken and written English, English articles often appear as marked features (Master, 2002; Myers, 1992). The learnability challenges L2 learners of English face with articles are often explained from a cross-linguistic perspective (Celce-Murcia & Larsen-Freeman, 1999). One concern obviously is that not every language has an article system as English does. Many L2 learners of English come from a Native Language (L1) background with no article system such as Chinese, Japanese, Korean, and Russian. Another concern is that among the languages that do have an article system, not every language organizes articles in the same way that English does (e.g. Samoan).

L2 learners of English whose L1 lacks an article system tend to find the acquisition of the English article system even more difficult than those learners whose L1 has an article system (Jarvis, 1999; Myers, 1992). For example, Korean does not have an article system like English does. In English, the number distinction (singular versus plural) and the distinction between new and shared information of a noun phrase (NP) are often marked by articles. In Korean however, noun phrases often appear in the bare form without any functional markers to indicate these features (Ree, 1975). The lack of an article system in Korean is apparent in simple sentences as in (1).

- (1) Na-nun ecey tosekwan-eyse chayk-ul pilli-ess-ta.  
I-TOP yesterday library-from book-OCC check out-PAST-DEC  
“I checked out a/the book (or books) from a/the library (or libraries).”

The noun, *chayk* (‘book’) in (1) can be interpreted either as an indefinite singular (*a book*) or a definite singular (*the book*), or it can be interpreted either as an indefinite plural (*books*) or a definite plural (*the books*). The noun, *tosekwan* (library) can also take interpretations of an indefinite singular, a definite singular, an indefinite plural, or a definite plural. In English, a specific definite entity whose referent is known both to the speaker and hearer is always overtly marked either with the definite article *the* or other grammatical aspects, such as possessives. In Korean, on the other hand, the definiteness of an entity is often understood implicitly from the context. Consequently, for adult Korean L2 learners of English, using an article before a noun phrase is a challenging task, and using the appropriate article in the given context is even more difficult. As a result, L2 English article acquisition studies have emphasized the omission of obligatory articles as one of the most commonly occurring errors among Korean L2 learners of English (Myers, 1992).

The English article system may not appear to be complex on the surface. Its form is relatively simple compared to other grammatical aspects, such as the various forms of verb tenses and complex structures such as relative clauses. However, one of the main reasons for the developmental delay in the L2 English article system appears to be the lack of direct form-function correlation in article use

---

\*The research reported here is based upon the first author’s M.A. thesis (Kim, 2006), which was directed by the second author, Usha Lakshmanan, who created the design of the experimental items used in the study.

(Master, 2002). As much as article choices are rule-based to L2 learners of English, they are also semantic-based (Donnellan, 1966) and pragmatic-based (Kripke, 1977; Ludlow & Neale, 1991). The semantic complexity of English articles along with the discourse context which the learners have to rely on make it exceptionally difficult for L2 learners to master the seemingly simple use of articles.

## 2. Article Semantics

### 2.1 Semantics of English Articles

Cross-linguistically, articles encode semantic distinctions such as *definiteness* and *specificity* (Ionin, Ko, & Wexler, 2004). The notion of *definiteness* refers to the state of knowledge shared between the speaker and hearer (or writer and reader). The notion of *specificity* refers to the state of knowledge known to the speaker (writer) only. The definitions of *definiteness* and *specificity*, proposed by Ionin et al., are provided in (2). The present study adopts these definitions of *definiteness* and *specificity*.

#### (2) *Definiteness* and *Specificity*: Informal definitions

If a Determiner Phrase (DP) of the form [D NP] is . . .

- a. [+definite], then the speaker and hearer presuppose the existence of a unique individual in the set denoted by the NP.
- b. [+specific], then the speaker intends to refer to a unique individual in the set denoted by the NP and considers this individual to possess some noteworthy property.

(Ionin, Ko, & Wexler, 2004, p. 5)

Gernsbacher and Robertson (2001) note that the English definite article *the* indicates that the referent is already shared between the speaker and hearer (Bock, 1977; Grieve, 1973; Haviland & Clark, 1974; Harris, 1974; Osgood, 1971). Irwin, Bock, and Stanovich (1982) point out that “the definite article *the* marks old, given, or presupposed information, while the indefinite article *a* marks new or asserted information” (p. 308). de Villers (1974) suggests that the definite article *the* signals referential coherence. On the other hand, the use of the indefinite article *a* in discourse informs the comprehender that a new entity is introduced (Murphy, 1984). In Standard English, article choices depend on the notion of *definiteness*, not *specificity*. For example, in a [+definite] context where the speaker and hearer have shared knowledge of an entity being referred to, the definite article *the* is used. In a [-definite] context in which no such shared knowledge exists between the speaker and hearer, indefinite articles (*a*, *an*, or *some*) are used. Sentences in (3) illustrate how the feature of *definiteness* plays a role in determining article choices in Standard English.

- (3) a. I saw a car.
- b. The car was running a red light.

The choice of the indefinite article *a* in (3a) indicates that there is no shared knowledge of the car between the speaker and hearer. It is a first mention of the car by the speaker, and therefore the sentence attributes the feature of [-definite]. In (3b) the speaker is referring to the same car that she just mentioned in (3a), and she uses the definite article *the*. With the choice of the definite article *the* in (3b), the hearer realizes that *the car* is the same one as in the previous utterance. As shown in (3), in Standard English, the choice of the definite article *the* indicates the semantic condition of [+definite], and the choice of an indefinite article indicates a [-definite] context. Let us now consider the role of the semantic feature *specificity* in article choice in English. First consider the sentences in (4).

- (4) a. I want to watch a movie.
- b. Any recent movie will be fine.

The choice of the indefinite article *a* in (4a) shows that it is a first mention of the movie by the speaker. There is no shared knowledge of the movie between the speaker and hearer, and thus the context is [-definite]. Further, the sentence in (4b) shows that the speaker does not have a particular movie that she/he is referring to, and thus the context is [-specific]. The semantic condition of the DP (*a movie*) in

(4a) is therefore [-definite, -specific]. The indefinite article *a* in (4a) receives a quantificational interpretation (Fodor & Sag, 1982); the speaker is not referring to a particular movie that he/she has in mind, but is merely asserting *one* movie out of a group of movies that is available. Consider now the sentences in (5).

- (5) a. I want to buy a book.  
 b. It's on the best-seller list.

The use of the indefinite article *a* in (5a) indicates that it is a first mention of the book by the speaker in the discourse, and thus the context is [-definite]. This time, however, the sentence (5b) shows that the speaker has a particular book in mind that he/she is referring to (i.e., the book which is on the best-seller list). The semantic context of the DP in (5a) is therefore [-definite, +specific]. The indefinite article *a* in (5a) receives a referential interpretation (Fodor & Sag, 1982).

As shown in (4) and (5), the feature of *specificity* does not play a role in article choice in English. Whether a noun phrase is [-specific] as in (4) or [+specific] as in (5), as long as it is [-definite], the indefinite article *a* is used. If the context is [+definite], then the definite article *the* is used.

## 2.2 Semantics of Articles in the Samoan Language

Recent studies have revealed cross-linguistic variation in article choice. For example, unlike in English, articles in the Samoan language are associated with the feature of *specificity*. There are two articles in Samoan: *le* (or *l*) and *se* (or *s*). Mosel and Hovdhaugen (1992) illustrate that the Samoan article *le/l* is used in [+specific] contexts, and *se/s* is used in [-specific] contexts. Unlike English, *definiteness* does not play a role in article choice in Samoan. Sentences in (6) illustrate the use of the Samoan article *le/l*.

- (6) Use of the Samoan article *le/l*  
 a. [-definite, +specific] context: *le/l*

'O le ulugāli'i, fānau l=a lā tama 'o le teine 'o Sina  
 PRES ART couple give birth ART=Poss 3.du. child PRES ART girl PRES Sina  
 "There was a couple who had a child, a girl called Sina."

(Mosel & Hovdhaugen, 1992, p. 259, ex. 6.37)

- b. [+definite, +specific] context: *le/l*

Māsani 'o le tamāloa e usua'i=ina lava ia....  
 used PRES ART man GENR get up early=ES EMPH 3sg  
 'ae nonofo 'o le fafine ma l=a=na tama i le fale  
 but stay(pl.) PRES ART woman and ART=POSS=3.sg child LD ART house  
 "It was the man's practice to get up early and... while the woman stayed at home with her child."

(Mosel & Hovdhaugen, 1992, p. 259, ex. 6.38)

In (6a), the speaker is telling a story about a couple with a child with whom he/she is acquainted. Since it is a first mention of the family, the context is [-definite], and there exists a particular family that the speaker is referring to, thus the context is also [+specific]. Notice that in the [-definite, +specific] context, the article *le* is used. In (6b) the speaker continues the story of the family, which indicates that the context has become [+definite]. Notice that the article choice remains the same; the same article *le* is used for the [+definite, +specific] context. The sentences in (6) show that regardless of the definiteness of the context, the article *le/l* is used if the context is [+specific].

Consider now the use of another Samoan article *se/s* in the examples provided in (7).

- (7) Use of the Samoan articles *se/s*  
 a. [-definite, -specific] context: *se*

'Au=mai se niu!  
 take=DIR ART(nsp.sg.) coconut  
 "Bring me a coconut [no matter which one]!"

(Mosel & Hovdhaugen, 1992, p. 261, ex. 6.46)

b. [+definite, -specific] context: *se*

*Alu i se tou aiga e moe. Pe se tama a ai!*  
 go LD ART(nsp.sg.) 2.pl. family GENR sleep. Q ART(nsp.sg.) boy POSS who

“Go to your family – whoever that may be – and sleep! [I wonder] whose boy you might be!”

[said to a boy who is selling necklaces at night in front of a hotel]

(Mosel & Hovdhaugen, 1992, p. 262, ex. 6.53)

In (7a), the speaker is asking for a coconut, but he/she is not referring to a particular one. The context is [-definite, -specific], and the article *se* is used. In (7b), there exists one particular family to which the boy belongs, and that makes the context [+definite], but the speaker does not necessarily know which family that is, and this makes the context [-specific]. Overall the context in (7b) is [+definite, -specific]. Regardless of the feature of *definiteness*, as long as the context is [-specific], the article *se* is used. The choice of the Samoan article *le/l* and *se/s* is based upon *specificity*.

### 2.3 Article Choice Parameter

Ionin, Ko, and Wexler (2004) speculated that if article acquisition is constrained by Universal Grammar (UG), article choices must be derived from parameter settings. As a result, they proposed a semantic parameter - that is, *Article Choice Parameter*. As described above, in English, articles are organized in line with *definiteness*, and in Samoan, articles are organized on the basis of *specificity*. Based on these cross-linguistic distinctions in article choice, Ionin et al. proposed two options for the Article Choice Parameter: (i) the *definiteness* setting and (ii) the *specificity* setting. There are languages, such as English, in which article choice is constrained by the *definiteness* setting. Other languages, such as Samoan, organize articles according to the *specificity* setting.

In Ionin, Ko, and Wexler (2004), Korean and Russian L2 English learners participated in a forced-choice article elicitation task and a written production task. The results revealed the overuse of the definite article *the* with specific indefinites, [-definite, +specific]. The authors concluded that L2 learners optionally associate the definite article *the* either with *definiteness* or *specificity*, that is they fluctuate between the *definiteness* parameter setting and the *specificity* parameter setting.

## 3. Research Questions:

This study attempts to verify the *specificity* hypothesis (Ionin, et al. 2004), which predicts that Korean L2 learners of English will fluctuate between the two settings of the Article Choice Parameter (i.e., *definiteness* setting and *specificity* setting). It represents a first attempt to connect the following two areas: (i) the semantic interpretation of English articles by L2 learners of English and (ii) real-time online sentence processing. In doing so, a real-time, self-paced online reading experiment and an offline semantic acceptability rating experiment were carried out. The main research questions addressed in the study are as follows:

1. Do Korean L2 learners of English distinguish the English definite article *the* and the indefinite article *a* on the basis of *specificity* rather than *definiteness*, as the theory claims, in a real-time online reading experiment?
2. Do Korean L2 learners of English distinguish the English definite article *the* and the indefinite article *a* on the basis of *specificity* rather than *definiteness*, as the theory claims, in an offline semantic acceptability rating experiment?

## 4. Methodology

### 4.1 Participants

The participants of the study consisted mainly of two groups: an experimental group consisting of adult native speakers of Korean, who are second language learners of English, and a control group of adult native speakers of English. The L2 experimental group was divided into two groups of Advanced and Intermediate based on their English proficiency levels, as determined by a cloze pretest. Both L2

groups consisted of nine participants. The Korean participants ranged in age from 18 to 47 ( $M = 29.21$ ) including 10 male and 8 female participants. The English control consisted of 14 participants (6 males and 8 females) ranging in age from 18 to 29 ( $M = 22.5$ ).

#### 4.2 Materials and Procedures

The instruments used in this study consisted of the following: (i) a written questionnaire, (ii) a self-paced, real time online reading experiment, (iii) an offline semantic acceptability rating experiment, (iv) a cloze pretest, and (v) an article insertion pretest. A questionnaire, written in English, was administered to obtain participants' demographic information. The instrument was adapted from the questionnaire developed by Usha Lakshmanan and a group of her students (Ji-Eun Kim, Susilowaty Margono, and Yuko Yamamori) in *Linguistics 542: Universal Grammar and Second Language Acquisition* in Spring 2004.

The cloze pretest was designed primarily to determine the Korean subjects' English proficiency levels, but was administered to the native speakers of English as well. The participants were asked to fill in systematically deleted blanks (every eighth word was deleted) in a passage, taken from a visitor's guide to Korea (Storey & Crowther, 1995). The instrument contained a total of 40 blanks to fill in, and the subjects were given 30 minutes to complete the exercise. The instrument was originally developed by James M. Ranalli (University of Birmingham) for his study, entitled "Comparing scoring procedures on a cloze test" conducted in November, 2002. It was minimally modified and implemented in the present study.

The article insertion pretest was administered to ensure that the Korean L2 learners of English have the English article morphemes in their interlanguage. The data from the main experiments of the study (i.e., online and offline experiments) would not be valid if participants did not have any knowledge of articles. Similar to the cloze pretest, the article insertion test was also text-based. The participants were given a short passage in which all the obligatory articles were deleted. Of a total of 13 blanks that required an English article, six represented the indefinite article *a* and seven represented the definite article *the*. There were three additional blanks representing words other than articles as distracters. The participants were asked to fill in the blanks with appropriate words. However, unlike the cloze pretest, the participants were not forced to fill in every blank. If they thought that no word was necessary in the blank, they were asked to fill in the blank with  $\emptyset$ . Also important to note, the instructions did not provide any articles for the participants to choose from, unlike many article elicitation instruments. If the participants were asked to fill in the blanks with articles, they could have merely inserted an article in the blank without having any knowledge of articles at all.

#### 4.3 Experiment 1: Self-paced, Real-time Online Reading Experiment

The word-by-word, self-paced online moving window reading task, called *Linger* (Rohde, 2001), was controlled by custom software on Mac computers. Each participant was seated in front of a computer monitor, and he/she was asked to read sentences presented on the screen. The stimulus sentences, as well as distracter sentences were presented in pairs. At the beginning, participants saw two lines as shown below presented on the computer screen.

---

The participants were instructed that the two lines represented the sentences spoken by one person, engaged in conversation with a friend, who she was meeting after a long time. The first line represented the first sentence in the pair, and the second sentence represented the second sentence. Words in every sentence appeared one by one by pressing the space bar. With every press of the space bar, a new word appeared, and the old one disappeared. *Linger* recorded the time between each press of the space bar (i.e., the time spent reading each word) in milliseconds. When the participants finished reading the last word of the second sentence, a question appeared on a new screen asking whether or not the second sentence in the pair made sense in combination with the first sentence. The participants were instructed to press the "F" key on the keyboard for YES and the "J" key for NO. The participants were instructed to read the sentences as naturally as possible, and to answer the questions as quickly as possible.

#### 4.4 Experimental Items

The experimental items comprised two distinctive semantic conditions: [-definite, -specific] and [-definite, +specific]. Each condition contained 20 experimental items. Each experimental item consisted of a pair of sentences. In each experimental item, the first sentence contained a singular noun phrase with an English article: either the indefinite articles *a(an)* or the definite article *the*. Of the 20 items in each condition, ten items contained the indefinite article *a*, and the remaining ten contained the definite article *the*. The sentences in these two sets were identical except for the article used in the first sentence. The second sentence in each pair provided the target semantic condition the researcher intended to manipulate: either [-definite, -specific] or [-definite, +specific]. Since both conditions encode the [-definite] context, this feature will be omitted hereafter for ease of presentation. Sample experimental items in the [-specific] and the [+specific] conditions are presented in (8) and (9) respectively.

(8) Sample experimental items in the [-definite, -specific] condition

a. [-definite, -specific *a*] item

I'm looking for a hotel. → contains the indefinite article *a*  
 Any cheap hotel is fine. → provides the [-definite, -specific] condition

b. [-definite, -specific *the*] item

I'm looking for the hotel. → contains the definite article *the*  
 Any cheap hotel is fine. → provides the [-definite, -specific] condition

(9) Sample experimental items in the [-definite, +specific] condition

a. [-definite, +specific *a*] item

I met an actor. → contains the indefinite article *a*  
 You'll never guess which actor I met. → provides [-definite, +specific] condition

b. [-definite, +specific *the*] item

I met the actor. → contains the definite article *the*  
 You'll never guess which actor I met. → provides [-definite, +specific] condition

#### 4.5 Predictions

Ionin, Ko, and Wexler (2004) predict that unlike adult native speakers of English, Korean L2 learners of English initially distinguish English articles on the basis of the *specificity* parameter setting rather than the *definiteness* parameter setting. In other words, Korean L2 learners of English will interpret the English indefinite article *a* as [-specific] rather than [-definite], and the definite article *the* as [+specific] in stead of [+definite]. If Ionin et al.'s *specificity* theory is valid, when the Korean participants read the first sentence with the indefinite article *a* in the [-specific] condition (8-a), they will build a [ $\pm$ definite, -specific] semantic context. Regardless of the semantic feature of *definiteness*, the article *a* is associated with a [-specific] interpretation. As they read the second sentence in the [-specific *a*] items, their expectations of a [-specific] context would be confirmed. Therefore the second sentence should be processed by the Korean L2 learners without processing interruption.

On the other hand, when they are reading the first sentence of the item with the definite article *the*, the Korean L2 learners of English would expect to see a [ $\pm$ definite, +specific] context. The English definite article *the* is the '[+specific] article' to the learners. The use of the article *the* in the first sentence leads the readers to believe that the speaker has a unique entity (NP) in mind and is referring to it. However, as they process the second sentence, they would notice that the speaker does not have a particular entity in mind that he/she is referring to (e.g., *Any cheap hotel is fine.*). The context provided by the second sentence is still [-specific]. The mismatch between the readers' expectation and the actual semantic condition provided by the second sentence would consequently slow down the reading time of the second sentence.

As far as the native speakers of English are concerned, it was predicted that they would read the second sentence in the [-specific *a*] items faster compared to the second sentence in the [-specific *the*] items. In the [-specific] condition, the use of the indefinite article *a* is appropriate, and the use of the definite article *the* is inappropriate. The inappropriate use of the definite article *the* in this condition would slow down the reading time of the second sentence for the native speakers of English. To

summarize, it was predicted that native speakers of English and the Korean L2 learners would have the same patterns in reading time in the [-specific] condition, but for different reasons as described above. The reading time of the second sentence in the [-specific *a*] items would be faster compared to the reading time of the second sentence in the [-specific *the*] items in each participant group.

In the [+specific] condition, it was predicted that the Korean L2 learners of English would read the second sentence in the [+specific *the*] items faster than the second sentence in the [+specific *a*] items. When they read the first sentence with the indefinite article *a*, they would expect to see a [±definite, -specific] context. Unlike their expectation, the second sentence provides a [+specific] context. The L2 learners would notice the mismatch between the use of the indefinite article *a* in the first sentence and the [+specific] semantic condition provided by the second sentence. This will slow down the processing time of the second sentence. On the other hand, when they read the first sentence which contains the definite article *the*, they would associate the article with a [+specific] context, and their expectations would be confirmed by the second sentence.

For the native speakers of English, the patterns in the reading time in the [+specific] condition would be the opposite of the experimental groups' reading time patterns. Again in this condition the native speakers of English would find the use of the definite article *the* inappropriate. Thus the reading time of the second sentence in the [+specific *the*] items compared to the reading time of the second sentence in the [+specific *a*] items would be longer.

#### 4.6 Distracter Items

Since the stimulus items contained pairs of sentences that are semantically compatible as well as incompatible according to the article choices, distracter items were created in the same manner. There were a total of 35 distracter items included in the online experiment, and they were presented in pairs as well. The same Yes/No comprehension question was presented after each distracter item as in the case of the stimulus items. The distracter items were created based on three conditions. The first set of distracter items consisted of 15 pairs of sentences, and the first and second sentences in each pair were semantically compatible with one another. The second distracter group also consisted of 15 pairs of sentences, but the two sentences in each item were semantically incompatible. The last group of distracters included five pairs of sentences in which the words in the second sentence were jumbled. This particular group of items was created to ensure that the participants were indeed reading the items, and not merely pressing the spacebar without reading the words. The Yes/No questions following the jumbled-up items required the 'No' response and participants who selected the 'YES' response more than three times were excluded from the data analysis.

#### 4.7 Experiment 2: Offline Acceptability Rating Task

An offline acceptability rating experiment was carried out to examine to what extent the participants accept or reject the stimulus items. The participants were asked to rate, on a four point Likert-type scale, the semantic acceptability of the first sentence in the stimulus items (i.e. the one with an English article). The offline task included the same experimental items used in the online task: 20 pairs of sentences in the [-specific] condition and 20 pairs of sentences in the [+specific] condition. In order to distract the participants from becoming aware of the grammatical focus of the study, whole new sets of distracter and jumbled-up items were created. Sample experimental stimuli in the [-specific] and [+specific] conditions used in the offline semantic acceptability task are shown in (10) and (11) respectively.

(10) Sample experimental items in the [-specific] condition used in the offline task:

- a. "I'm looking for *a* hotel. Any cheap hotel is fine."
- b. "I'm looking for *the* hotel. Any cheap hotel is fine."

(11) Sample experimental items in the [+specific] condition used in the offline task:

- a. "I met *an* actor. You'll never guess which actor I met."
- b. "I met *the* actor. You'll never guess which actor I met."

Each stimulus was followed by a question, asking if sentence underlined in the pair of sentences makes sense together with the sentence that was not underlined. The scale provided ranged from 0 to 3 points (0: 'Doesn't make sense at all'; 1: 'Somewhat doesn't make sense'; 2: 'Somewhat makes sense'; 3: 'Absolutely makes sense').

In the offline task, it was hypothesized that the native speakers of English would accept the sentences with the indefinite article *a* (grammatical sentences), but reject the sentences with the definite article *the* (ungrammatical sentences) in both [-specific] and [+specific] conditions. For the Korean L2 learners of English, it was hypothesized that they would accept the [-specific *a*] items, and reject the [-specific *the*] items. They would however reject the [+specific *a*] items even though they were grammatical, and accept the ungrammatical [+specific *the*] items. These predictions were based on the theory that the Korean L2 learners of English associate the definite article *the* with [+specific] rather than [+definite]. The use of the indefinite article *a* in the [+specific] context would consequently be considered as unacceptable by the learners.

## 5. Procedures for Data Analysis

The data management and analysis for the online and offline tasks was performed using SPSS 13.0 (2005). All statistical comparisons were made using a paired samples *t*-test within each subject group. Since the aim of the study was to compare the responses in the two sub-conditions in each semantic condition (i.e., [±specific *a*] and [±specific *the*]), no other statistical procedure was deemed necessary. For instance, comparing the values among the native, advanced L2, and intermediate L2 groups using an ANOVA test was not relevant because it is generally assumed that non-native speakers of English would read the sentences at a slower rate than native speakers do. Another important reason for not comparing the values (e.g., reading time and semantic acceptability rating scores) between the subject groups was that the most relevant information sought was whether these values within the subject groups were consistent. The relevant comparisons were not between the native, advanced, or intermediate L2 groups but among each subject group as a whole.

The alpha level (i.e. significance criterion) used for all the statistical tests was .025. Since the data from the same individuals were compared twice for the [-specific] and [+specific] conditions on the online and offline tasks, the original alpha level .05 was divided by 2, yielding an alpha level of .025.

The raw reading time of individual words in every sentence was recorded by *Linger*, and the data was then analyzed by importing the *Linger* data files into SPSS data analysis software. The crucial value for statistical analysis was the amount of time spent reading the entire second sentence (i.e., *Total* reading time) in each experimental item rather than the reading time of a particular word or region of the sentence. A particular word or region in the second sentence could not be selected for statistical comparison because each experimental item was comprised of an unequal number of words. For instance, the second sentence of every item in the [+specific] condition was designed with a different structure, such as “*Can you guess which NP...?*” and “*You won't believe which NP...*” It was necessary to vary the structure of the second sentence since each condition had to contain exactly the same items except for the article in the first sentence.

In order to compare the reading time of the second sentence, the following procedures were undertaken. First, the reading time of the individual words in the second sentence in each of the stimulus items was summed for the *Total* reading time. The *Total* reading time of every word in the second sentence for each of the ten items in each sub-category (i.e., [-specific *a*], [-specific *the*], [+specific *a*], and [+specific *the*]) was computed. Then the *Mean* reading time of the *Total* reading time for each subject group was obtained. The *Mean* reading time was obtained by dividing the *Total* reading time of the ten second sentences obtained from all subjects in each group by the number of subjects. This calculation was carried out in SPSS. The term *Mean* reading time (or *reading time* for ease of presentation) used hereafter refers to the *Mean* of the *Total* reading time of the second sentence in all ten items in each sub-condition. The reading time of the second sentence in the [±specific *a*] items was compared to the reading time of the second sentence in the [±specific *the*] items using a paired samples *t*-test within each subject group.



## 6. Results

### 6.1 Online reading time in the [-definite, -specific] condition

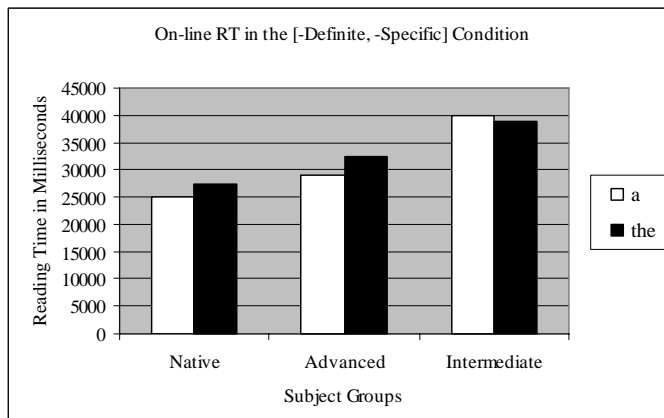
Table 1. reports the Mean Reading Times (RTs) and the Standard Deviations of the [-definite, -specific] condition for the three groups (native speakers, Intermediate learners, and Advanced learners).

Table 1.  
Mean Reading Times (RTs) in Milliseconds and Standard Deviations of the [-specific] Condition

Group	N	Sub-condition	Mean RT	Std. Deviation	Mean Difference	p-value
Native	14	[-specific <i>a</i> ]	24904.57	8184.282	-2567.571	.056
		[-specific <i>the</i> ]	27472.14	8095.860		
Advanced	9	[-specific <i>a</i> ]	29067.11	9114.444	-3291.444	.257
		[-specific <i>the</i> ]	32358.56	15830.634		
Intermediate	9	[-specific <i>a</i> ]	40007.11	16450.673	1131.333	.663
		[-specific <i>the</i> ]	38875.78	14783.055		

The *Mean* difference value in Table 1 was obtained by subtracting the *Mean* reading time of the [-specific *the*] items from the *Mean* reading time of the [-specific *a*] items. The negative value in the *Mean* difference, therefore, indicates that the *Mean* reading time of the [-specific *the*] items was larger than the *Mean* reading time of the [-specific *a*] items. The positive value in the *Mean* difference, on the other hand, indicates that the *Mean* reading time of the [-specific *the*] items was smaller than the *Mean* reading time of the [-specific *a*] items. Figure 1. provides a visual representation of the patterns in the *Mean* reading times observed in each sub-condition for the three participant groups.

Figure 1. Mean Reading Time of the second sentence in the [-definite, -specific] Condition



Note: RT refers to (Mean) Reading Time.

As shown in Table 1 and Figure 1, the native speaker control group took longer on average to read the second sentence in [-specific *the*] items than the second sentence in the [-specific *a*] items. Although the patterning of the reading times was in the predicted direction, the difference between the *Mean* reading times (-2567.571) was not significant ( $t(13) = -2.095, p = .056$ ). The advanced Korean L2 group, like the native speaker control group, also took longer to read the second sentence in [-specific *the*] items than the second sentence in the [-specific *a*] items. Although the patterning was in the predicted direction, the difference between the *Mean* RTs for these two sub-conditions (-3291.444 milliseconds) was not significant ( $t(8) = -1.219, p = .257$ ). The intermediate L2 group, unlike the other two groups, took longer to read the second sentence in the [-specific *a*] items than the second sentence in the [-specific *the*] items. The *Mean* difference in reading times, which was 1131.333 milliseconds, was not found to be significant ( $t(8) = .453, p = .663$ ). In sum, statistically significant differences between the reading time of the [-definite, -specific *a*] items and the [-definite, -specific *the*] items were not observed for any of the three subject groups.

## 6.2 Online reading time in the [-definite, +specific] condition

A paired-samples *t*-test was conducted to evaluate the hypothesis that there was a difference in the *Mean* reading time of the second sentence in the [+specific *a*] items and the *Mean* reading time of the second sentence in the [+specific *the*] items in each participant group. Table 2 compares the *Mean* reading times and the Standard Deviations of the [+specific] items within each subject group.

Table 2.

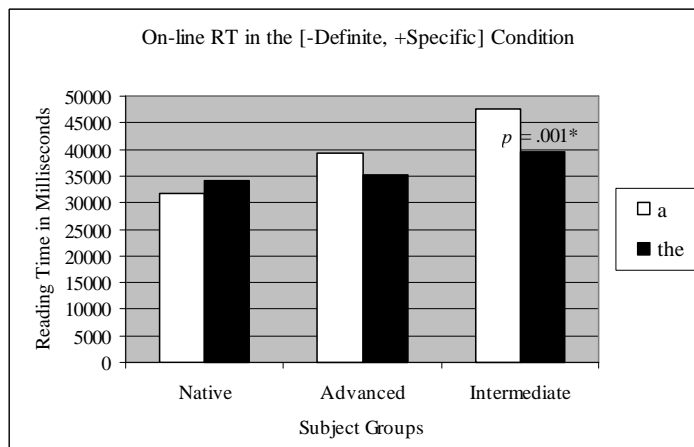
Mean Reading Times (RTs) and Standard Deviations of the [-definite, +specific] Condition

Group	N	Sub- condition	Mean RT	Std. Deviation	Mean Difference	p-value
Native	14	[+specific <i>a</i> ]	31885.93	7804.676	-2329.500	.089
		[+specific <i>the</i> ]	34215.43	10429.325		
Advanced	9	[+specific <i>a</i> ]	39280.11	14043.968	4070.444	.170
		[+specific <i>the</i> ]	35209.67	10632.032		
Intermediate	9	[+specific <i>a</i> ]	47822.56	18540.279	8173.778	.001*
		[+specific <i>the</i> ]	39648.78	17366.710		

Note: Asterisk (\*) indicates the statistical significance at .025 alpha level.

Figure 2 provides a visual representation of the patterns in the *Mean* reading time of the second sentence in the [+specific *a*] items and the *Mean* reading time of the second sentence in the [+specific *the*] items for each of the three groups.

Figure 2. Mean Reading Time of the second sentence in the [-definite, +specific] Condition



The asterisk (\*) indicates the statistical significance at .025 alpha level.

As shown in Table 2 and Figure 2, the native-speaker control group read the second sentence in the [+specific *a*] items faster than the second sentence in the [+specific *the*] items. The *Mean* difference between the two reading times was -2329.500 milliseconds. Although the trends in the reading times were in the predicted direction, the difference between the *Mean* reading times was not significant,  $t(13) = -1.837, p = .089$ . The advanced Korean L2 group, unlike the native speaker control group, read the second sentence in the [+specific *the*] items faster than the second sentence in the [+specific *a*] items. The *Mean* difference value was 4070.444 milliseconds. Even though this outcome is consistent with our prediction for the [-definite, +specific] condition, the difference between the *Mean* reading times was not significant ( $t(8) = 1.508, p = .170$ ). As in the case of the advanced Korean L2 group, the intermediate Korean L2 group read the second sentence in the [+specific *the*] items faster than the second sentence in the [+specific *a*] items. The *Mean* difference value was 8173.778 milliseconds. Unlike in the case of the advanced L2 group, the difference between the *Mean* reading times was found to be significant, thus supporting our prediction in relation to the [-definite, +specific] sub-condition. In sum, only the intermediate L2 learners, but not the advanced L2 learners, processed the second sentence in the [+specific *the*] items *significantly* faster than the second sentence in the [+specific *a*] items. The native control and the advanced L2 groups showed *no significant* difference

between the reading time of the [+specific *a*] items and the [+specific *the*] items, although the trends in the patterning of the reading times, in each case, was in the predicted direction.

### 6.3 Offline ratings in the [-definite, -specific] condition

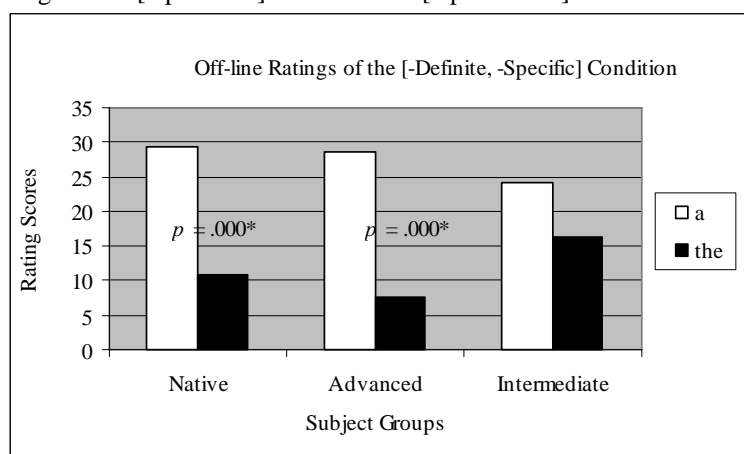
A paired-samples *t*-test was conducted to evaluate the hypothesis that there was a difference in the *Mean* rating of the [-specific *a*] items and the *Mean* rating of the [-specific *the*] items. The *Mean* ratings and Standard Deviations obtained from the paired samples *t*-tests for the [-specific] condition are presented in Table 3. Figure 3, provides a visual representation of the rating patterns of the [-specific *a*] items and the [-specific *the*] items.

Table 3. Mean Ratings and Standard Deviations of the [-definite, -specific] Condition

Group	N	Sub- condition	Mean Rating	Std. Deviation	Mean Difference	p-value
Native	14	[-specific <i>a</i> ]	29.2143	2.22498	18.35714	.000*
		[-specific <i>the</i> ]	10.8571	7.16677		
Advanced	9	[-specific <i>a</i> ]	28.6700	2.179	21.00000	.000*
		[-specific <i>the</i> ]	7.6667	10.44031		
Intermediate	9	[-specific <i>a</i> ]	24.1111	9.06152	7.77778	.060
		[-specific <i>the</i> ]	16.3333	11.68332		

Note: Asterisks (\*) indicate the statistical significance at .025 alpha level.

Figure 3. Mean Ratings of the [-specific *a*] items and the [-specific *the*] items



The asterisk (\*) indicates the statistical significance at .025 alpha level.

As can be seen from Table 3 and Figure 3, the native speakers behaved as predicted and rated the [-specific *a*] items more highly than the [-specific *the*] items. The *Mean* difference between the two ratings was 18.35714, which was significant ( $t(13) = 9.339, p = .000$ ). The advanced L2 group, similar to the native speakers, behaved as predicted, and rated the [-specific *a*] items more highly compared to the [-specific *the*] items. The *Mean* difference between the two ratings was 21, which was significant ( $t(8) = 6.126, p = .000$ ). The Intermediate L2 group also rated the [-specific *a*] items higher than the [-specific *the*] items. Although the trend was in the predicted direction, the *Mean* difference between the two ratings (7.77778) was not found to be significant. Furthermore, as shown in Figure 3, they demonstrated a higher acceptance rating for the [-specific *the*] items than did the other two groups. In sum, the predictions in relation to the [-definite, -specific] condition in the offline task were supported only in the case of the native speaker control group and the advanced L2 group but not in the case of the intermediate group.

#### 6.4 Offline ratings in the [-definite, +specific] condition

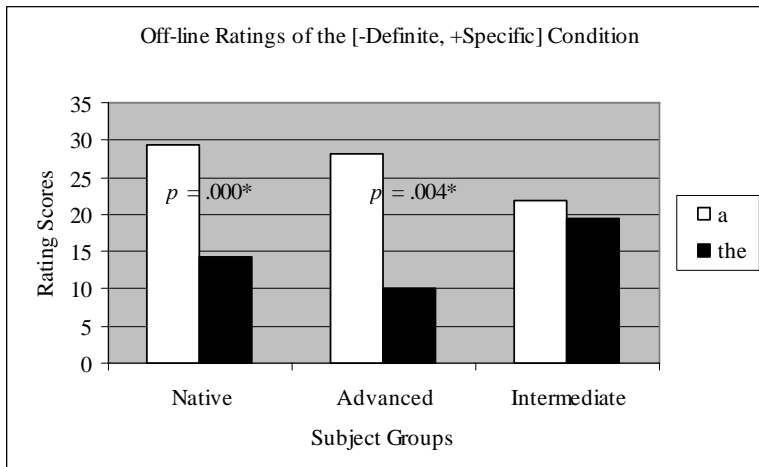
A paired-samples *t*-test was conducted to evaluate the hypothesis that there was a difference in the *Mean* ratings between the [+specific *a*] items and [+specific *the*] items within each participant group. Table 4 presents the *Mean* ratings and Standard Deviations of the [+specific *a*] and the [+specific *the*] items. Figure 4 provides a visual representation of the rating patterns of the [+ specific *a*] items and the [+ specific *the*] items.

Table 4. Mean Ratings and Standard Deviations of the [-definite, +specific] Condition

Group	N	Sub- condition	Mean Rating	Std. Deviation	Mean Difference	p-value
Native	14	[+specific <i>a</i> ]	29.3571	1.39268	14.92857	.000*
		[+specific <i>the</i> ]	14.4286	5.27278		
Advanced	9	[+specific <i>a</i> ]	28.2222	3.19287	18.11111	.004*
		[+specific <i>the</i> ]	10.1111	11.65595		
Intermediate	9	[+specific <i>a</i> ]	21.7778	10.98610	2.22222	.470
		[+specific <i>the</i> ]	19.5556	12.00116		

Note: Asterisks (\*) indicate the statistical significance at .025 alpha level.

Figure 4. Mean Ratings of the [+specific *a*] items and the [+specific *the*] items



The asterisk (\*) indicates the statistical significance at .025 alpha level.

As can be seen from Table 4 and Figure 4, the native speakers behaved as predicted and rated the [+specific *a*] items more highly compared to the [+specific *the*] items. The *Mean* difference between the two ratings was 14.92857 which was significant ( $t(13) = 11.005, p = .000$ ). The advanced L2 group, similar to the native speakers, but contrary to what was predicted, rated the [+specific *a*] items more highly compared to the [+specific *the*] items. The *Mean* difference between the two ratings was 18.11111, which was significant ( $t(8) = 3.999, p = .004$ ). The results for the native speaker control group and the advanced L2 group indicates that they tended to accept the [+specific *a*] items and reject the [+specific *the*] items. The intermediate L2 group did not appear to distinguish between the acceptability of the [+specific *a*] items and the [+specific *the*] items. The *Mean* difference between the two ratings was only 2.22222, and it was not significant ( $t(8) = .758, p = .470$ ). Contrary to what was predicted, the intermediate L2 learners did not prefer the use of the definite article *the* over the use of the indefinite article *a* in the [-definite, +specific] condition in the offline task.

## 7. Discussion and Conclusion

For the online task, it was predicted that in the [-definite, -specific] condition, the native speakers of English would read the second sentence in the “*a*-items” faster than the second sentence in the “*the*-items”. The same prediction was made for the advanced and the intermediate L2 groups, but for different reasons as noted previously. However, contrary to what was predicted, no statistically

significant difference was observed between the reading time of the [-specific *a*] items and the reading time of the [-specific *the*] items for the three participant groups. These results indicate that the different uses of the articles in the first sentence did not have a significant effect on the reading time of the second sentence in the [-specific] condition.

In the [-definite, +specific] condition, it was predicted that the native controls would read the second sentence in the [+specific *the*] items slower than the second sentence in the [+specific *a*] items. It was predicted, however, that the advanced and intermediate L2 groups would read the second sentence in the [+specific *the*] items faster than the second sentence in the [+specific *a*] items. The native control and the advanced L2 group did not exhibit significant differences between the reading times of the [+specific *a*] items and the [+specific *the*] items. However, compelling evidence was obtained that the intermediate L2 learners of English read the second sentence in the [+specific *the*] items significantly faster than the second sentence in the [+specific *a*] items. This suggests that the intermediate L2 learners of English found the use of the definite article *the* in the first sentence appropriate in the [+specific] condition, but the use of the indefinite article *a(n)* inappropriate. This further suggests that the learners associated the English definite article *the* with [+specific] rather than [+definite], and the indefinite article *a(n)* with [-specific] rather than [-definite]. This finding is consistent with the findings of Ionin et al's. (2004) study, based on a forced-choice production task, that intermediate level Korean L2 learners of English tended to associate the definite article *the* with [+specific] rather than [+definite].

The results from the offline semantic acceptability rating task revealed that in both [-definite, -specific] and [-definite, +specific] conditions, the native control group and the advanced L2 group strongly accepted the first sentence which contained the indefinite article *a(n)*, but rejected the first sentence which contained the definite article *the*. This means that they found the use of the indefinite article *a(n)* in both conditions appropriate, and the use of the definite article *the* inappropriate. This further indicates that they associated the English indefinite article *a* with [-definite] and the definite article *the* with [+definite] regardless of whether the context is [-specific] or [+specific]. Both the native speakers of English and the advanced L2 learners adhered to the *Definiteness* setting of the Article Choice Parameter in the offline semantic acceptability task.

Unlike the native control group and the advanced L2 group, in the offline task, the intermediate L2 learners did not distinguish between *a* and *the* in both semantic conditions (i.e., [-specific] and [+specific]). The results from the online reading task, however, strongly suggest that they associated the English definite article *the* with [+specific] rather than [+definite]. Recall that they processed the second sentence in the [+specific *the*] items significantly faster than the second sentence in the [+specific *a*] items. The intermediate L2 learners seem to initially distinguish between the definite article *the* and indefinite article *a* on the basis of *specificity* as was evident in their performance on the online task. However, when they were given more time and greater opportunities to reflect upon their responses in the offline task, the learners might have been able to consider other options in the article parameter settings, which consequently led them to fluctuate between the two parameter settings, as was evident from their offline results.

The native control and the advanced L2 group showed a strong adherence to the *Definiteness* parameter setting in the offline task. Yet this was not the case in the online task. One possible explanation for this is that in the online task, the participants might have been reconstructing (i.e., substituting) what they actually read with what they thought they had read in terms of articles. By the time they were reading the second sentence, they might have already mentally replaced the incorrect article with the correct one (i.e., replaced *the* with *a*) in both semantic conditions.

The other explanation for the inconsistent performance of the native control and the advanced L2 group is that they might have been fluctuating between the two settings of the Article Choice Parameter. The advanced L2 learners' parameter settings may have to do with the level of consciousness. In the online task, they might have been tapping more unconscious processing, whereas in the offline task, they might have been more conscious of the forms (the articles in this case). Furthermore, recall that in the online task, each word that was previously viewed disappeared from the screen with each press of the spacebar (that is, they could not look back at what they had read but had to rely on their memory instead). We speculate that their interpretation of articles might stem from both *specificity* and *definiteness*. Since they had access to both parameter settings in real-time sentence processing, their reading time of the second sentence in the [+specific] condition would not have differed depending on the choice of the article in the first sentence. Both the indefinite article *a* and the

definite article *the* might have been acceptable to them, resulting in similar reading times for the second sentence in the two sub-conditions. However, in the offline task, the learners might have been able to correct their initial hypothesis concerning the article choice parameter setting. As they were able to read the items as many times as they needed in the offline task, it might have allowed them to change their initial hypothesis. However in real-time processing, their mind was moving along with the words as they appeared on the screen, and they could not go back, check the article, and confirm or revise their hypothesis. Their responses were therefore based on their initial, intuitive hypothesis. Based on the results from both the intermediate L2 group and the advanced L2 group, we conclude that the Korean L2 learners' interpretation of English articles may initially stem from the *specificity* setting of the Article Choice Parameter and that it continues to persist at more advanced levels, even after the emergence of the *definiteness* setting.

We also speculate that the native speakers of English could have been fluctuating between the two settings of the Article Choice Parameter. There is evidence that child L1 learners of English initially organize articles on the basis of *specificity* rather than *definiteness* (Brown, 1973; Cziko, 1986). Like adult L2 learners of English, they initially tend to overuse the definite article *the* in [+specific] contexts. It is only at a later period that they begin to associate the definite article with [+definite] contexts. The results of our online task for the adult native speaker group suggest that the *specificity* parameter setting may still be available to native speakers during real-time processing.

Based on these findings, the following conclusion can be drawn from this study. Korean L2 learners of English initially tend to organize English articles on the basis of *specificity* rather than *definiteness*. However, as their proficiency level increases, they begin fluctuating between the two article choice parameter settings: the *specificity* parameter setting and the *definiteness* parameter setting.

## References

- Bock, J. K. (1977). The effect of a pragmatic presupposition of syntactic structure in question answering. *Journal of verbal learning and verbal behavior*, 16, 723-734.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL teacher's course*. Boston: Heinle & Heinle.
- Cziko, G. A. (1986). Testing the language bioprogram hypothesis: A review of children's acquisition of articles. *Language*, 62, 878-898.
- de Villiers, P. A. (1974). Imagery and theme in recall of connected discourse. *Journal of Experimental Psychology*, 103, 263-268.
- Donnellan, K. (1966). Reference and definite descriptions. *The Philosophical Review*, 75, 281-304.
- Fodor, J. D., & Sag, I. A. (1982). Referential and quantificational indefinites. *Linguistics and Philosophy*, 5, 355-398.
- Gernsbacher, M. A., & Robertson, R. R. W. (2001). The definite article as a cue to map thematic information. In M. M. Louwse & W. Van Peer, *Thematic: Interdisciplinary studies*. Amsterdam/Philadelphia: John Benjamins.
- Grieve, R. (1973). Definiteness in discourse. *Language and Speech*, 16, 365-372.
- Harris, R. J. (1974). Effects of nonlinguistic knowledge on language production. *Journal of Verbal Learning and Verbal Behavior*, 3, 303-310.
- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of verbal learning and verbal behavior*, 13, 512-521.
- Heim, I. (1991). Articles and definiteness. [published in German as "Artikel und Definitheit"], in A. V. Stechow & D. Wunderlich (Eds.), *Semantics: An International Handbook of Contemporary Research*. De Gruyter: Berlin.
- Ionin, T., Ko, H., & Wexler, K. (2004). Article semantics in L2-acquisition: The role of specificity. *Language Acquisition*, 12(1), 3-69.
- Irwin, D. E., Bock, J. K., & Stanovich, K. (1982). Effects of information structure cues on visual word processing. *Journal of verbal learning and verbal behavior*, 21, 307-325.
- Jarvis, S. (1999, September). Universal intra- and interlingual processes in English article acquisition. Paper Presented at SLRF '99, Minneapolis/St. Paul, MN.
- Kim, K. (2006). *The interpretation and processing of English articles by native speakers of Korean*. M. A. thesis, Southern Illinois University, Carbondale.

- Kripke, S. (1977). Speaker's reference and semantic reference. In P. A. French, T. e. Uehling, Jr., & H. Wettstein (Eds.), *Midwest studies in philosophy, volume II: Studies in the philosophy of language*. (255-276). University of Minnesota. Morris, MN
- Ludlow, P., & Neale, S. (1991). In defense of Russell. *Linguistics and philosophy*, 14, 171–202.
- Lyons, C. (1999). *Definiteness*. Cambridge University Press: Cambridge, England.
- Master, P. (2002). Information structure and English article pedagogy. *System*, 30, 331-348.
- Mosel, U., & Hovdhaugen, E. (1992). *Samoan reference grammar*. Oslo: Scandinavian University Press.
- Myers, S. (1992). In search of the genuine article: A cross-linguistic investigation of the development of the English article system in the written compositions of adult ESL students. Unpublished Ph.D. dissertation, The Florida State University.
- Osgood, C. E. (1971). Where do sentences come from? In D. D. Steinberg & L. A. Jakobovits (Eds.), *Semantics* (pp. 497–529). Cambridge University Press: Cambridge, England.
- Poesio, M., & Vieira, R. (1998). A corpus-based investigation of definite description use. *Computational Linguistics*, 24(2) 183-216.
- Ranalli, J. M. (2002). Comparing scoring procedures on a cloze test. Retrieved on November 23, 2005, from <http://www.cels.bham.ac.uk/resources/essays/Ranalli5.pdf>
- Ree, J. J. (1975). Demonstratives and numbers in Korean. In H. Sohn (Ed.), *The Korean language* (pp. 33-46). University of Hawaii.
- Rohde, D. (2001). *Linger*. Retrieved February 4, 2004, from <http://tedlab.mit.edu/~dr/Linger/>
- Storey, R., & Crowther, G. (1995). *Korea*. Hawthorn, Lonely Planet Publications.

# Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)

edited by Roumyana Slabakova, Jason Rothman, Paula Kempchinsky, and Elena Gavrusseva

Cascadilla Proceedings Project Somerville, MA 2008

## Copyright information

Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)  
© 2008 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-422-5 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.  
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

## Ordering information

Orders for the library binding edition are handled by Cascadilla Press.  
To place an order, go to [www.lingref.com](http://www.lingref.com) or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA  
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: [sales@cascadilla.com](mailto:sales@cascadilla.com)

## Web access and citation information

This entire proceedings can also be viewed on the web at [www.lingref.com](http://www.lingref.com). Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Kim, Kyoungsook and Usha Lakshmanan. 2008. L2 Article Semantics and Second Language Processing. In *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)*, ed. Roumyana Slabakova et al., 103-117. Somerville, MA: Cascadilla Proceedings Project.

or:

Kim, Kyoungsook and Usha Lakshmanan. 2008. L2 Article Semantics and Second Language Processing. In *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)*, ed. Roumyana Slabakova et al., 103-117. Somerville, MA: Cascadilla Proceedings Project. [www.lingref.com](http://www.lingref.com), document #1630.