A COMPREHENSIVE STUDY OF THE USE OF ENGLISH ARTICLES

BY KOREAN L2 SPEAKERS OF ENGLISH:

SPEAKING, WRITING, AND METALINGUISTIC AWARENESS

BY

JUYEON LEE

A dissertation submitted to

The Graduate School of Education

Rutgers, The State University of New Jersey.

in partial fulfillment of the requirements

for the degree

Doctor of Education

Approved by

_____

Dr. Nydia Flores, Chair

_____

Dr. Mary Curran, Committee

_____

Dr. Liliana Sánchez, Committee

New Brunswick, New Jersey

January, 2013

ABSTRACT

A COMPREHENSIVE STUDY OF THE USE OF ENGLISH ARTICLES
BY KOREAN L2 SPEAKERS OF ENGLISH:
SPEAKING, WRITING, AND METALINGUISTIC AWARENESS

BY

Juyeon Lee

In the field of second language acquisition (SLA), it has been documented that the mastery of the English article system by English language learners (ELLs) is a challenging task. To understand the ways in which ELLs acquire the use of articles and the underlying causes of the difficulties associated with this particular feature, a comprehensive study was conducted on Korean L2 speakers with the purpose of understanding implicit and explicit knowledge of English articles. The study was informed by several theoretical frameworks in SLA and sociolinguistics, and thus employed a mixed-method approach to its analysis to address the following questions:

1. What patterns in the use of articles are evidenced in Korean L2 speakers of English at distinct levels of English language proficiency?
2. How are these patterns distinct in speaking and writing?
3. What kinds of metalinguistic awareness do Korean L2 speakers of English have at different English language proficiency levels?
4. How are hypotheses about the use of English articles of Korean speakers of English different from those of native English speakers?
5. How are Korean L2 English speakers' hypotheses about English articles operationalized in their speech?

Four tasks were performed by 30 Korean learners of English from three different proficiency levels: oral and written narratives, grammaticality judgment tasks, and think-aloud tasks. In addition, five native English speakers (NES) undertook the grammaticality judgment

task and the think-aloud task to serve as a comparative basis.

The study revealed that article use varied not only according to the external variable of language proficiency but also with respect to several linguistic variables such as NP contexts in which a noun phrase contained an adjective. In addition, the study found that articles were generally produced with higher rates of accuracy in writing rather than in speech. With regard to the explicit knowledge of English articles, the results of the think-aloud protocol suggest that ELLs may gradually acquire the manner in which nominal entities are constructed with regard to countability and definiteness and that learners are able to incrementally understand how the use of articles is conditioned by semantic, pragmatic, and contextual features. The study also addresses several pedagogical implications regarding how articles are understood by learners whose first languages do not have a similar article system and how article instruction needs to focus on the pragmatic, semantic, and contextual features.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter I. Introduction

1.1. Background

In the field of second language acquisition (SLA), it has well been documented that the mastery of the English article system by learners of English as a second language is a challenging task (Butler, 2002; Ekiert, 2010; Ionin, Ko, & Wexler, 2004; Liu & Gleason, 2002; Park, 2006; Parrish, 1987; Trenkic, 2007; Young, 1996; Zegarac, 2004, among others). As Master (1997) points out, the articles are among the most frequently used English words. Therefore, learners are likely to be exposed to abundant input that contains a variety of uses of articles. Nevertheless, successful acquisition of the article system by second language (L2) learners of English seems to be difficult, if not impossible (e.g., Ekiert, 2010; Huebner, 1979; Master, 1990; Mizuno, 1999; Park, 2006; Parrish, 1987; Whitman, 1974).

Factors affecting these difficulties may be many, but two major challenges are prominent in the literature. First, there seems to be a lack of one-form to one-meaning correspondence. As Young (1996) points out, the article system in English encodes "semantic notions of existence, reference, and attribution; discourse notions of anaphora and context; and syntactic notions of countability and number" (p. 135). English language learners may find it daunting to acquire this complex system that lacks a simple mapping between form and meaning.

The second major challenge is concerned with L2 learners whose native languages do not have a similar grammatical feature, including Korean, Japanese, Chinese and Russian (see Ekiert, 2010; Ionin et al., 2004; Robertson, 2000; Young, 1996, among others). Different languages may grammaticalize such distinctions as 'definiteness' and 'specificity' of noun referents in different ways, and learners find it difficult to understand the English article system

when their first languages (L1s) do not express semantic, discourse, and syntactic features of noun referents in a similar way as English does.

It is important that we understand the challenges that English language learners face in learning the English article system. This knowledge can contribute to developing effective ways to teach this feature and understand aspects of second language acquisition. A wealth of research on the topic of L2 acquisition of English articles has been conducted, but there still are missing pieces of the mosaic to understand the ways in which English language learners acquire the use of articles and the underlying causes of the difficulties associated with this particular feature of English.

This study aims to fill in some of those missing gaps by conducting a comprehensive study on how Korean L2 speakers understand and use English articles in view of implicit knowledge (i.e., knowledge *of* a language as the ability to use the language fluently and accurately) and explicit knowledge (i.e., knowledge *about* a language, as in knowing the rules for the use of articles). There are three objectives this study aims to achieve. The first objective is to investigate the implicit knowledge of English articles manifested in the oral and written production of articles by Korean learners of English. The second objective is to explore the explicit knowledge of English articles verbally expressed by the participants as they make article choices for the blanks provided in a think-aloud task. The final objective is to compare spontaneous use of articles in oral narratives with explicit hypotheses about articles obtained from the think-aloud protocol in order to investigate differences between the participants' implicit and explicit knowledge of English articles. Results of the quantitative and qualitative data analyses will be discussed from the perspective of conceptual changes for second language acquisition (SLA). Findings from this study will contribute to the field of English as a Second

Language (ESL) by providing ESL teachers with a better understanding of how to approach this aspect of English in their teaching.

1.2. The Focus of the Study

This section explains how the research questions for this study have emerged. Second language acquisition of English article has been studied from a multitude of angles and using a variety of research methods. Several strands of research interests have developed over time with regard to L2 acquisition of English articles and these interests have motivated the current study.

First, I recognized a need for a comprehensive study on the use and acquisition of English articles by Korean learners of English. Previously, several findings have been reported in the literature with regard to article use by participants from various L1 backgrounds. For instance, Zegarac (2004) argued that non-suppliance of an article is the default for learners whose L1s do not have a similar system. Thomas (1989) and Parrish (1987) showed that the definite article *the* is acquired before the indefinite article *a*. A number of researchers documented that learners go through a period of over-generalization of the definite article (Ionin et al., 2004; Liu & Gleason, 2002; Parrish, 1987; Thomas, 1989; Zegarac, 2004). Also, Ionin et al. (2004) further argue that learners display a fluctuation between the definiteness setting and the specificity setting in their article choices. A number of learner populations have been studied with regard to L2 acquisition of English articles, but there has been relatively little research on English language learners from Korean L1 background. As a native speaker of Korean, I had struggled to understand appropriate usage and functions of English articles as Korean does not have an article system similar to English. Thus, I developed interest in investigating how Korean learners of English acquire and use articles at different proficiency levels and examining whether research findings on learners of

English from other article-less L1 backgrounds are also evidenced in the Korean population.

The second topic that was investigated in this study is concerned with the written production of L2 articles compared to the oral production of articles. There have been several studies in which oral language was collected for analysis of article usage (Huebner, 1979; Parrish, 1987; Robertson, 2000; Tarone & Parrish, 1988; Thomas, 1989; Trenkic, 2007; White, 2003; Zdorenko & Paradis, 2008). However, little has been documented about how learners of English use articles in their writing. For instance, Ellis, Sheen, Murakami, & Takashima (2008) examined article usage of English language learners in their written narratives, but the focus of the study was to test the effectiveness of written corrective feedback. Lee's (1997) study was based on free essay writing of Korean university students, but it only examined instances in which articles were omitted in obligatory contexts.

In conjunction with the first and the second topics, this study compared article use in oral and written narratives. The purpose of the comparisons was to investigate what aspects of article use would benefit most from the absence of time constraints. In spontaneous oral production, L2 learners are under pressure to attend to accuracy and fluency simultaneously. Without time constraints, learners may be able to direct their attention to grammatical accuracy with regard to article use. Thus, it can be assumed that the overall accuracy in the use of articles would be better in written production, and this study aimed to explore in more detail how the different discourses of speaking and writing would affect patterns of article use among Korean learners of English at different levels of proficiency.

Third, this study was designed to explore explicit knowledge of Korean learners of English with regard to the use of English articles. Here, *explicit* knowledge refers to declarative knowledge that can be brought into awareness and is available for verbal report. Previously,

Butler (2002) interviewed Japanese learners of English after the participants completed a fill-in-the-blank task and asked for reasons behind their article choices. This study was informed by Butler's (2002) findings and investigated whether Korean learners of English at various proficiency levels possess similar kinds of metalinguistic knowledge of the English article system as reported in Butler (2002).

A major difference between Butler (2002) and this study is that the participants in this study provided reasons for their article use as they were making article choices in the think-aloud protocol, whereas the participants in Butler (2002) were interviewed for metalinguistic hypotheses of English articles after they completed the article choice task. However, as Ericsson and Simon (1993, as cited in Bowles & Leow, 2005) point out, retrospective protocols, like the protocol used in Butler (2002), need to be used with caution because it is possible that the information obtained at the time of the verbal report may be different from the information participants actually used while performing the task. In contrast, as Akakura (2009) explains, a think-aloud protocol is an effective measure of explicit knowledge that allows direct access to thought processes and minimizes forgetting and post-hoc rationalization of decisions. Thus, this study utilized a concurrent think-aloud protocol.

Fourth, the various kinds of metalinguistic understanding of the English article system by Korean L2 speakers of English were compared with those provided by a control group of native English speakers. The comparison between hypotheses that L2 speakers have formed with regard to articles and those presented by native speakers was expected to provide insights on how to bridge the gap between L2 learners of English and native English speakers.

Finally, in this study, hypotheses generated by the participants were compared with their oral and written language in a qualitative manner to capture their meta-cognitive understanding

of English articles. Butler (2002) investigated metalinguistic knowledge of Japanese learners of English through a fill-in-the-article task and qualitative interviews with individual participants. Following the report of the results, Butler (2002) raised the question of the role of implicit knowledge and its relation to explicit knowledge. In connection with this question, the scholar commented that "it is not clear to what extent the learners actually access metalinguistic knowledge for article use when they engage in regular production activities, such as free speaking and free writing" (p. 475). This study intended to advance our understanding on the issue of the degree to which the metalinguistic hypotheses reported by L2 speakers of English are reflected in their actual speech.

With the development of the above-mentioned research interests, this study was guided by the following research questions.

1. What patterns in the use of articles are evidenced in Korean L2 speakers of English at distinct levels of English language proficiency?

2. How are these patterns distinct in speaking and writing?

3. What kinds of metalinguistic awareness do Korean L2 speakers of English have at different English language proficiency levels?

4. How are hypotheses about the use of English articles of Korean L2 speakers of English different from those of native English speakers?

5. How are Korean L2 English speakers' hypotheses about English articles operationalized in their speech?

1.3. Outline of the Dissertation

The remainder of this dissertation is organized as follows. In Chapter II, several strands

of research motivating the current study are reviewed. The chapter begins with an overview of a cognitive approach to second language learning, and a brief description of Cognitive Grammar is presented. Next, theories of the English article system are reviewed with the focus on the concepts of definiteness, followed by a discussion of the concept of the so-called 'zero article'. The following section focuses on empirical research on L2 acquisition of the English article system is reviewed. Finally, the research questions of the study are presented again, together with the significance of this study.

Chapter III presents the design and methodological procedures of the study. The chapter begins with a description of methodological concerns in the study of English articles. Next, the participants are introduced with regard to recruitment procedures and English language proficiency testing. This section is followed by a description of the study design. Then the data gathering process is described, including data collection instruments and procedures. The chapter ends with a description of the data analysis procedures employed.

Chapter IV reports the results of the quantitative analyses of oral and written production of articles. In order to address the first research question, the distinct patterns of article use by the participants in their oral narratives are analyzed and discussed with respect to their overall English language proficiency levels. This segment is followed by the analyses of their written production of English articles. Following the analyses of the patterns of article use that emerged from the oral and written corpus, the data are compared between the two discourses and the proficiency levels with the purpose of addressing the second research question which focuses on similarities and differences in article use in the written and spoken discourses.

In Chapter V, results of the grammaticality judgment task and the think-aloud task are reported with the purpose of investigating explicit knowledge of articles exhibited by the learners

as opposed to the native speaking control group. First, performance results of the grammaticality judgment task are reported for each proficiency group. The grammaticality judgment task was administered as a supplementary measure of explicit knowledge of articles. Next, the kinds of metalinguistic knowledge that participants manifested regarding their knowledge of English articles are presented for each proficiency level to address the third research question that focused on the kinds of metalinguistic awareness of the Korean learners of English. Then, in response to the fourth research question, explicit hypotheses of English articles generated by the participants during the think-aloud task are compared with those of native English speakers. Finally, three case studies are reported in which the implicit knowledge of English articles manifested in the oral narratives is compared with the explicit knowledge about English articles obtained from the think-aloud protocol. These case studies serve to address the last research question that aims to investigate how explicit understanding of articles is operationalized in the oral production of articles.

Chapter VI summarizes and discusses major findings of this current study. First, patterns of article use that emerged in the oral and written corpus are discussed in light of previous research findings. Then, metalinguistic hypotheses of article use developed and expressed by the participants are discussed and compared with those expressed by the NES participants. These results are considered as they relate to the literature. This section is followed by a discussion of the interplay between implicit and explicit knowledge of English articles. This chapter concludes by discussing theoretical and pedagogical implications of the current study.

Finally, Chapter VII discusses limitations of the study and proposes directions for future research.

Chapter II. Review of the Literature

This chapter is devoted to reviewing several strands of research motivating the current study. First, a description of the cognitive approaches to second language acquisition is provided as an underlying theoretical framework of this study. Next, the English article system is discussed from structural, semantic, and pragmatic perspectives. This section is followed by an overview of previous research on the acquisition of English articles by L2 learners. Specifically, the review of relevant previous studies will focus on a range of factors directly related to acquisitional difficulty of the English article system. Finally, the research questions of the study are re-introduced.

2.1. Cognitive Approach to Second Language Acquisition

As a theoretical framework, the current study has been guided by cognitive linguistics. This section provides a brief review of a cognitive approach to grammar, which centers on the semantic and communicative nature of language rather than its formal properties. Within this framework, grammar is assumed to reflect how language users carve up events and experiences. Thus, a cognitive view of grammar posits meaning as the dynamic activity of human minds interacting with their environment. Among several theories in the cognitive linguistics tradition, this current study was mainly informed by the tenets of Cognitive Grammar proposed by Langacker (2008b).

Langacker (2008b) explains that there are three basic features of Cognitive Grammar: the centrality of meaning, the meaningfulness of grammar, and its usage-based nature. Cognitive Grammar posits that grammar is meaningful and elements of grammar have meanings in their own right. Grammar allows us to construct and symbolize the more elaborate meanings of

complex expressions.

There are two particular aspects of Cognitive Grammar that have relevance to how English articles are used: (1) distinctions between count and mass nouns, and (2) nominal grounding. The first aspect is related to the usage of the indefinite article *a* as this article form is applied to a singular countable referent. According to Langacker (2008b), a count noun is construed as "being bounded within the immediate scope in the domain of instantiation" (p. 132). Conversely, the profile of a mass noun is not construed as being bounded in this manner. The immediate scope can be understood as the locus of attention. Langacker explains that the count/mass distinction hinges on whether bounding occurs within the immediate scope or not. While the boundary of a count noun falls within the immediate scope, a mass noun does not invoke a boundary within its immediate scope of attention. In Cognitive Grammar, bounding does not necessarily coincide with the physical nature of nominals. For instance, *wood* and *board* may have the same physical attributes, but *board* is conceptualized by the speakers of English as a bounded entity while *wood* is not.

The second aspect of Cognitive Grammar concerning articles is nominal grounding. Grounding, in general, is described as a semantic function, which is "an aspect of conceptual organization by which an expression qualifies as a nominal or a finite clause" (p. 272). In English, there are certain elements that serve a grounding function: the articles (*the*, *a*), demonstratives (*this*, *that*, *these*, *those*), and certain quantifiers (*all*, *most*, *some*, *no*, *every*, *each*, *any*). They have their specific grammatical properties, but they all play a role in signaling out a nominal referent. The current study did not adopt the manner in which articles are theorized in Cognitive Grammar, but was informed by the notion that grammatical features are meaningful and that grammatical distinctions are concerned with how speakers of a language interpret and

conceptualize the events and experiences in the world. In other words, the current study posits that articles in English make distinct semantic contributions to overall utterances in a discourse and that articles hinge on how speakers of English understand nominal entities in their cognitive domain and how they express them in communication.

Cognitive Grammar also bears important implications for second language acquisition and research. Langacker (2008b) claims that grammar is a mental reflection of the way we conceptualize the world through all our experiences. According to this view, it can be posited that speakers of different languages may have different ways of conceptualizing the world around them. Thus, important differences exist in how different languages carve up events and experiences in cognitively, socially, and culturally meaningful ways. In this sense, Dirven (1989) argued that discovering and instructing the conceptualizations of the form is facilitative of L2 acquisition, particularly in instances in which the L2 differs from the L1.

Since languages differ with regard to the aspects of reality that they give prominence to, a certain conceptual change must occur in the learner to construct a second language grammar. A notion related to conceptual restructuring and changes in L2 acquisition is the idea of conceptual transfer. While cross-linguistic transfer in second language learning is often discussed in the context of differences and similarities between the structural properties of the L1 and the target language, Jarvis and Pavlenko (2008) point out another kind of transfer: similarities and differences in conceptual categories corresponding to grammatical categories of the L1 and L2. They identified eight domains of reference that allow speakers to talk about themselves and their surroundings: objects, emotions, personhood, gender, number, time, space, and motion. When these grammaticalized concepts are different between the L1 and L2, learners may rely on L1-mediated concepts, resulting in non-target-like patterns of language use. This kind of cross-

linguistic influence has been termed by Jarvis and Pavlenko (2008) as conceptual transfer.

With regard to the notion of conceptual transfer, Han (2008) related the L2 morphology problem to the L1 based conceptual system. The scholar made a case for English plural –*s* interfacing with the conceptual system. He argued that persistent non-target-like plural marking in the L2 production of L1 Korean learners stems from the influence of an L1-based semantic and conceptual system. That is, the manner in which noun numbers are marked in Korean has impact on the acquisition of discourse-distributional properties of the target form for plural marking in English. For pedagogical effectiveness, the author suggested that the plural –*s* must be semantically and conceptually related to the abstract notion on number, which, in turn, intersects with other abstract notions such as (in)definiteness. Second, the author theorized that the underlying concepts available for L2 learners may eventually render the plural – *s* unlearnable for some learners. These findings have implications for the current study. Similar to the plural –*s* morphology, Korean does not have an article system similar to English. As such, conceptual re-organization of the semantics of nominal definiteness is needed for Korea learners to be able to acquire the proper use of articles.

Several studies on L2 article acquisition have recently employed cognitive frameworks in investigating second language use and pedagogy of English articles. For instance, based on a cognitive view of second language acquisition, Akakura (2009) studied the effect of explicit instruction on English articles on measures of both implicit and explicit knowledge of articles. The study showed that explicit instruction plays a facilitative role in the development of L2 language knowledge on articles in terms of both implicit and explicit knowledge. Pedagogically, these results suggest that explicit instruction on English articles based on a cognitive grammar may be beneficial to English L2 users, particularly in promoting the development of implicit

knowledge of article use.

In another study, Ekiert (2010) applied the framework of Cognitive Grammar in analyzing article use by L1 Slavic speakers of English. The scholar suggested that, for adult language learners, the construction of an L2 grammar often requires a revision of relevant language-mediated concepts. Relating the idea of conceptual restructuring in L2 acquisition to the patterns of article use observed in the study, Ekiert claimed that the notion of perspective-taking presents a considerable obstacle for target-like use of articles by L1 Slavic speakers. In other words, the notion of familiarity to the hearer may be one of the potential variables responsible for the difficulty that English articles pose for L1 Slavic learners. In conclusion, Ekiert explains that the findings lend support to the view that full acquisition of grammatical elements, such as English articles, depends on conceptual restructuring from L1 to L2.

White (2010) applied a cognitive approach to grammar in developing a conceptual framework for English articles. The design of the framework was based on the assumption that meanings of articles can be adequately accounted for in a unified manner by using the notion of conceptualizations. In this framework, all uses of *the* map to a single schematic image, all uses of *a* and unstressed *some* map to a second schematic image, and all uses of the zero article map to a third schematic image. When this conceptual framework of the articles was presented to five MA TESOL students with L1s of Korean, Thai, and Chinese through a series of training sessions, the study found that explanations for article use were more unified across individual article uses after exposure to the framework.

This current study has also been shaped by Cognitive Grammar. In doing so, it assumes that English articles should be understood as meaningful elements that contribute to particular conceptualizations of the world around the language users rather than some morphological forms

that are combined with nominals based on certain mechanical rules. This study also recognizes

that article choices in English are dependent on how the speaker wants to present nominal

referents and the speaker's assessment of what is assumed to be shared with the hearer in their

immediate discourse contexts.

This section presented a brief description of Cognitive Grammar and its implications for

this current study as an over-arching theoretical framework. In the next section, the literature on

the English article system, a target L2 structure for this study, is reviewed from a multitude of

perspectives.

2.2. Understanding of English Articles

The grammatical structure under investigation in this study is the English article system.

Therefore, it is necessary to review this particular aspect of English grammar. In this section, a

brief sketch of the origins of English articles is provided. Then, various theoretical approaches to

the study of English articles are presented from structural, semantic, and pragmatic perspectives.

The section concludes with the discussion of the notion of the zero article. In doing so,

Berezowski's (2009) view on the rejection of the zero article is introduced as it informed this

study.

2.2.1. Etymological Aspect of English Articles.

Historically, the definite article and the indefinite article originated from different

sources. Hewson (1972) provides a detailed historical account of the article system. According to

Hewson, in Old English, some usages of the demonstrative pronouns can be translated by the

definite article in modern English, others by a demonstrative and still others by either the definite

article or a demonstrative. The shift from the original demonstrative to the modern article took place slowly over centuries. A key ingredient of the process is believed to be a gradual increase in the frequency of use of a grammatical form until it became obligatory for all nominals. This view was summarized in Christophersen (1939):

> In those Indo-European languages that possess a definite article it originated in a demonstrative pronoun. Like other innovations in language, the article started in a few isolated instances and gradually spread to all the cases where a common name is used in proper name capacity. It is clear, however, that in point of principle we cannot recognize it as an article until the development is completed and the word has become obligatory in all such cases. (Christophersen, 1939, p. 83)

Demonstrative expressions involve physical or verbal pointing and denote a distance from a referred entity to the interlocutors. In the gradual shift from a demonstrative pronoun to a definite article, *the* slowly lost the meaning of distal pointing and started to be used as an indication that the hearer can easily identify the referent denoted by the nominal in the immediate discourse context.

With regard to the indefinite article, the numeral *an* used in early Old English with the sense of *a certain* began to develop into an article during the Old English period. Hewson (1972) shows that *sum* (*some* in Modern English) was also used for individualizing indirect reference in Old English. Therefore, Hewson argues that in Old English, "there is no binary contrast of definite and indefinite, as in Modern English, but a contrast of definite and zero on the one hand, and on the other a contrast of *an* and *sum* for specific individualizing reference" (p. 21). By the 12th century, however, *an* took over the uses of *sum*. The indefinite article used in Chaucer in the 14th Century Middle English is similar to its use in Modern English (Hewson, 1972).

In sum, the definite article and the indefinite article in English are historically derived from different forms. The definite article *the* originated from demonstrative pronouns and the

indefinite article *a* from the numeral one. Both these articles have undergone a process of grammaticalization in which the use of these forms have expanded into other grammatical environments and become obligatory for nouns. In the next section, the syntactic representation of articles is discussed to understand the structural relationship between articles and nouns.

2.2.2. Structural Aspects of English Articles.

In the structural and formal study of language, articles are considered to belong to a category called Determiners. The Determiner category may include articles, demonstratives, quantifiers, possessive pronouns and so forth, which compete for a pre-nominal position. In a syntactic approach, therefore, formal characteristics of articles are often analyzed as subsumed by Determiners rather than treated in a separate way from other types of determiners.

The study of language focusing on the formal nature of language began to gain popularity in mid-20th century. In the Extended Standard Theory in the 1970s and early 1980s, a phrase was taken as a unit of analysis, and phrase structure rules, summarized in X-bar theory, were proposed as universal properties of human language. In its earlier versions of phrase structure rules, a Noun Phrase (NP) was considered to have a noun as the *Head* and take a determiner such as an article as an optional constituent in the *Specifier* position.

Figure 2. 1. The structure of a Noun Phrase
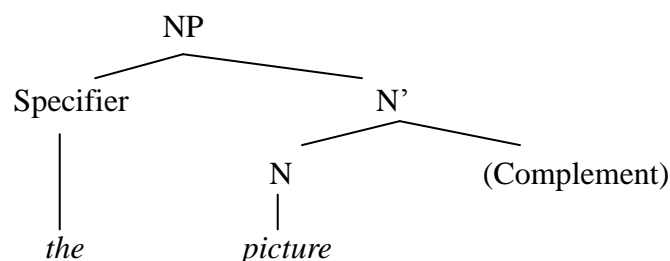
However, from the 1980s, the focus of linguistic research was shifted to finding evidence for universal features of human language. Summed as Universal Grammar, the Principles and Parameters model assumes universal formal representations of language. Under this framework, it has been noted that there are important internal parallels between clauses and noun phrases. Based on this observation, Abney (1987, cited in Bohnacker, 1997) and other scholars have suggested that both clauses and noun phrases are headed by functional categories. In this regard, it has also been argued that Determiner is taken to be the head of a functional projection, dubbed Determiner Phrase (DP) (Progovac, 1998). Then, an example of a DP structure is as follows. The article is the head of DP and NP is in the complement position.

Figure 2. 2. The structure of a Determiner Phrase
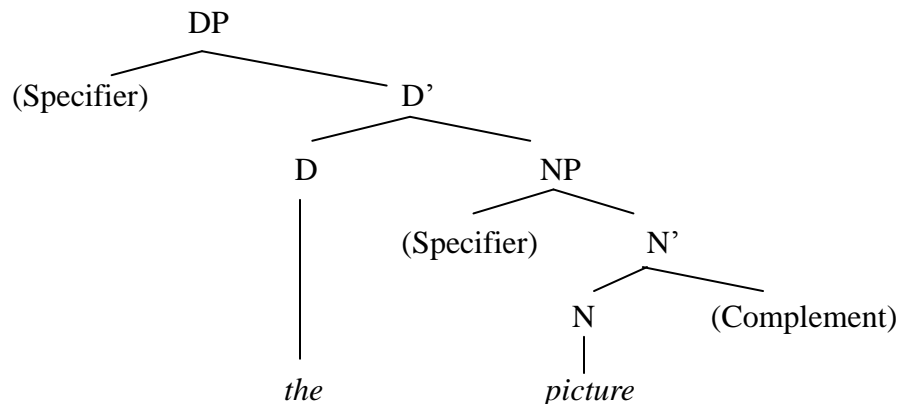
The DP hypothesis has been adopted by many generative linguists. It is generally assumed that the following elements constitute determiners and are thus generated under D: definite articles, indefinite articles, possessives, personal pronouns, demonstratives, and genitive or possessive *–s* (Bohnacker, 1997).

As discussed above, syntactic analyses of English articles help us understand how articles are structurally represented in relation to nouns in English. However, there are several

limitations to this approach in developing a comprehensive understanding of how the English

article system works. First, the primary focus of a formal analysis has been the investigation of

how different parts of a sentence are related structurally, which has not examined notional

relationship between parts of a sentence. Thus, particular grammatical and semantic functions

that articles serve are not dealt with separately in this framework. Articles are understood as a

type of Determiner along with other pre-nominal elements. Moreover, more recent linguistic

analyses have been focused on testing assumptions of Universal Grammar, and any attention to

articles in English was rather incidental. In order to understand how scholars have attempted to

formulate definitions of English articles, the following section examines the literature on the

analysis of the notion of definiteness.

2.2.3. Semantic Aspects of English Articles.

As implied by the names *definite article* and *indefinite article*, the notion of definiteness

is central to semantic analyses of the article system. This section discusses several scholars who

have made key contributions to the understanding of definiteness in English.

In the tradition of notional analyses of definiteness, the work of Christophersen (1939) is

considered a seminal piece that offers valuable insights concerning semantic distinctions of

articles in English. Christophersen considered *the* as an article of familiarity and argued that the

speaker can use the definite article when addressing the hearer if there is a mutual familiarity

with the referent. This is referred to as the familiarity theory.

> A condition of the use of *the* is that there is a basis of understanding between
> speaker and hearer. This basis comprises the subjects and things known by both
> parties, and the speaker as the active party must consequently adapt his language to
> the hearer's state of mind. If he wants to be understood it is important that he
> should not use words and phrases which the hearer is likely to misinterpret. (p. 28)

It is important to note that by referring to the notion of mutuality, Christophersen placed the perspective of the hearer at the forefront in his theorization of familiarity.

Christophersen's notion of familiarity can be compared to an analysis of the English articles by Jespersen's (1949). Jespersen's account is essentially a theory of "stages of familiarity" where familiarity is defined as "knowledge of what item of the class denoted by the word is meant in the case concerned" (p. 437). Stage I is complete unfamiliarity and corresponds to indefiniteness (e.g. a drink, milk). Stage II is near complete familiarity, in which the referent is to be found in the linguistic context or non-linguistic situation, and corresponds to the use of *the* (e.g. the house). In instances of near complete familiarity, the use of definite article is necessary to confirm the familiarity, as in the following (p. 417-148):

(1) Determination is given by the context:
Once there lived an old tailor in the village. *The tailor* was generally known in the village as the crook.

(2) Determination is given by the whole situation:
the government [of the time, of the country concerned]
Long live *the King*!
The school [that one to which he was sent] was an old and famous one
the sun [that one among many suns of which it is most natural to speak]
the devil

Stage III is complete familiarity, which makes the use of the article unnecessary. Proper names are examples of instances of complete familiarity.

However, it is not clear how to tell the difference between complete familiarity and near familiarity. For example, Jespersen's stages of familiarity cannot properly account for proper names that contain the definite article. Moreover, as Lyons (1999) pointed out, Jespersen's notion of familiarity is different from Christophersen in that Jespersen took the question of familiarity on the speaker's side while Christophersen conceived familiarity from the hearer's perspective.

Christophersen's (1939) original work has been built upon and expanded by other researchers as well. The concept of familiarity proposed by Christophersen (1939) was refined by Hawkins (1978) by introducing the notion of a shared set. According to Hawkins, the speaker performs the following acts when using *the*:

> He (a) introduces a referent (or referents) to the hearer; and (b) instructs the hearer to locate the referent in some shared set of objects…and he (c) refers to the totality of the objects or mass within the set which satisfy the referring expression. (p.167)

A shared set is established when both speaker and hearer share knowledge that a given referent is located in a given set. Examples of such shared sets, corresponding to usage types of the definite article, are:

> (a) anaphoric use: previous discourse between speaker and hearer
>   (e.g. Fred was discussing an interesting book in his class. I went to discuss *the book* with him afterwards.)
> (b) immediate situation and visible situation use;
>   (e.g. Pass me *the bucket* please.)
> (c) larger situation use: shared general knowledge
>   (e.g. I'll see you in *the pub* tonight.)
> (d) associates of the referent
>   (e.g. Fred has written a book. *The title* is 'Zen for Beginners'.)
>   (Examples provided by Chesterman, 1991, p. 18)

The use of the indefinite article, however, is neutral to the locatability of the referent. The referent with the indefinite article may be locatable in a pragmatic sense, but non-locatability is not a defining feature of indefinite reference.

Hawkins (1991) revised and updated his own original ideas (1978) on definiteness by using the theory of implicature (Grice, 1975, cited by Hawkins, 1991). Implicature is a meaning that is deducted from the form of an utterance. Messages are often not a direct composite of words in an utterance, but implicit messages need to be inferred by reading between the lines. There are two kinds of implicatures: conventional implicatures as a fixed element in the meaning

of an expression; and conversational implicatures, which may be cancelled by other pieces of information provided in the context. He argued that the implicature theory provides a theoretical framework for an integrated pragmatic-semantic-syntactic approach. Thus, the implicatures he proposed would "permit the indefinite article to maintain its simple logical translation as an existential quantifier" and would support the definition of "the 'uniqueness' of definite descriptions in a general and coherent way" (p. 406).

Hawkins (1991) also newly labeled a "shared set" as "pragmatic set" or "P-set" and redefined it in terms borrowed from Sperber and Wilson (1986) as a "mutual cognitive environment" (p. 414) – a set of facts manifest equally to the individuals concerned. In this framework, *the* entails existence and uniqueness, and carries a conventional implicature that there is some P-set accessible to the speaker and the hearer within which existence and uniqueness holds. *A* entails only existence and conventionally implicates non-uniqueness. Also, *a* can conversationally implicate P-membership provided the referent is non-unique in the P-set. Where the referent is unique in a P-set, *the* must be used. Thus, in the sentence:

(1)   I didn't buy *the house*, because *a window* was broken.

In the discourse context, there is a unique house both the speaker and the hearer have access to in their mind, thus allowing the use of *the*. In this example, *the house* is unique in the sense that the speaker and the hearer already have a shared understanding of a particular house that is relevant to this particular discourse context. On the other hand, there are many windows in this house, and one of them is broken. In this case, the use of *a* is appropriate due to its non-uniqueness in the P-set, in spite of its P-membership. Hawkins improved on Christopher's familiarity by taking into consideration the notions of conventional and conversational implicatures.

Chesterman (1991) embraced most of Hawkin's (1978) underlying assumptions

concerning locatability and inclusiveness in his comprehensive survey of the English article

system. Yet, he attempted to expand it to include the notion of extensivity, which reflects "the

difference between any surface article and no surface article" (p. 2). He adopted Guillaume's

(1919, cited in Chesterman, 1991) analysis on the distinction between any article and no article,

where "the zero article signifies that a noun is being used in its most abstract sense" (p. 26) and

adding an article presents it as a distinct entity or set of entities. In this sense, the notion of

extensivity is a distinction between "abstractness and generality" (p. 2). The notions of

locatability, inclusiveness and extensivity that Chesterman proposed as English article semantics

can be summarized as the following:

> "To distinguish these satisfactorily, the traditional opposition of definite vs. indefinite
> has been analyzed here as a composite of three more primitive semantic oppositions. One
> has to do with the relation of the NP to its context of use: locatability. Another is an
> opposition of quantity: all vs. not-all. Both these have been discussed at length in terms
> of Hawkin's location theory; yet if the description is also to account for the kinds of
> usage with 'no article', a third opposition must be incorporated, which, following
> Guillaume and Hewson, I have called limited vs. unlimited extensivity." (p. 88)

Chesterman argued that the combination of these oppositions allow more accurate descriptions of

various uses of the article system in English. He also offered useful insights on article usage:

First, when there is a conflict between a pragmatic interpretation and a strictly logical one with

regard to the understanding of locatability and inclusiveness by the speaker and the hearer,

pragmatics takes precedence. Second, article usage is connected with issues of quantity, and thus

the notions of reference and quantity seem closely intertwined. Finally, concepts such as

'generic', 'definite' and 'indefinite' are not unitary but more complex composites.

Lyons (1999) provided a detailed review of previous accounts of definiteness and found

that identifiability and inclusiveness are recurring concepts. Lyons argued that a distinction needs

to be made between grammatical definiteness and semantic/pragmatic definiteness. His proposal

is that "definiteness *stricto sensu* is not a semantic or pragmatic notion as assumed by almost all writers on the subject, but rather a grammatical category" (p. 274-275). In order to understand Lyons' proposal, it is important to understand the concept of grammaticalization, since definiteness is an instance of grammaticalization of some category of meaning.

Lyons (1999) held that the grammaticalization of definiteness may not be realized in the same way across languages. In fact, many languages completely lack overt definiteness marking, while some languages have definiteness in pronominal phrases but not in full, noun phrases. Lyons argued that lack of overt definiteness marking does not mean that the notion of definiteness does not exist in some languages. Rather, based on the separation between semantics and grammatical categories, it is argued that the semantic and pragmatic concept of definiteness is a universal concept regardless of its overt grammaticalization. Lyons used the idea of grammaticalization defined as "the representation by a grammatical form or forms of some concept of meaning" (p. 276) to account for cross-linguistic variations where some languages lack separate morphology to mark definiteness and maintain universality of the notion of definiteness.

Lyons proposed that in English, definiteness is the grammaticalization of identifiability of a referent, where the referent is uniquely identifiable to the speaker and the hearer in the given context. However, there would be other uses of definiteness unrelated to identifiability. Lyons explained that "when identifiability comes to be grammaticalized as definiteness, this category will go on to develop other uses" (p. 278). One such use is inclusiveness in Hawkin's (1978) term. His contribution to the debate on the notion of definiteness is the emphasis on how articles became grammaticalized in English while other languages did not experience a similar process. However, Lyons's (1999) discussions on the synchronic meaning of 'definiteness' are relatively

limited.

This section has reviewed how scholars have theorized the notion of definiteness in various ways to develop a unified account of the concept. Discussion of semantic attributes of definiteness is very important, but most researchers in this notional trend do not apply their theories in empirical studies. Similar to researchers who conduct structural analysis, the logical analysis mainly depends on constructed examples. Therefore, purely notional approaches to English articles are limited in the sense that such descriptions do not address how such semantic attributes of articles in English are operationalized in actual communication. Thus, we now turn to the discourse and pragmatic aspects of English articles in the next section.

2.2.4. Discourse Aspects of English Articles.

Within the context of a discourse that involves the speaker and the hearer, articles contribute to a successful conservation by providing information and clues that enable the construction and interpretation of the definiteness status of a referent. Therefore, in using articles, not only the speaker's but also the hearer's perspective is involved. For instance, Reid (1991) held that articles in English provide the hearer with additional information necessary for appropriate interpretation of a word in a given context. Specifically, *the* signals that the hearer has sufficient information to differentiate a referent, while *a* signals that the hearer does not have access to such information required for a complete identification.

With regard to the role of articles in a discourse, Trenkic (2000) held that general principles of goal-oriented behavior can account for the process of how the definiteness status of referents in a dialogue is established. The underlying assumption is the universal category of definiteness, articulated by Lyons (1999). Trenkic examined Serbian, a language without articles,

24

and English, and suggested that the task of a speaker in a language with articles is similar to the task of a hearer in a language without articles. In languages with articles, the speaker is responsible for determining the definiteness status of a referent and marking it grammatically for the hearer. In languages without articles, it is the hearer's responsibility to interpret the definiteness status of a referent, using various textual and contextual cues. In both cases, the goal is to achieve successful communication, and the speaker in languages with articles and the hearer in languages without articles make use of cognitive and linguistic strategies to achieve this goal.

Approaching the use of English articles in the context of a discourse from the framework of functional linguistics, Schiffrin (2006) carried out a functional analysis as she studied how speakers maintain and switch references in narratives. In English, using the definite and indefinite articles is a major way of structuring new and given information as well as maintaining referential cohesion in a discourse. In particular, Schiffrin (2006) was interested in investigating article repairs which is useful in addressing problems that arise when speakers and hearers work together to manage referrals.

According to Schiffrin, *the* is a formal indicator of definiteness, as well as pronouns, proper nouns and possessives, while *a* is a formal indicator of indefiniteness, together with quantifiers and numerals. According to Schiffrin, "the articles *the* and *a* establish a presumption of intersubjectivity that underlies how a referent is conceptualized by the speaker in his/her own mental model and how it is verbalized in a way that eases its representation in the hearer's mental model" (p. 70). Schiffrin maintained that on the one hand, articles indicate the speaker's level of specific knowledge about an entity. On the other hand, articles convey the information status of a referent to the hearer, which is an indication of the degree to which an upcoming referent is assumed to be recognizable in the context.

Findings of Schiffrin's analysis of article repairs show that some article switches reflect a functional convergence between *the* and *a*. Article switches refer to instances in which speakers start with one article and shortly switch to another for the same referent in the same discourse. Schiffrin argued that speakers may switch articles not only when they miscalculated the information status of a referent but also when they had trouble finding the fit between articles and the physical reality. As for article repetitions, the distribution of both definite and indefinite article repetitions generally follow their typical occurrence in sentence position. In summary, speakers constantly evaluate the information status of a referent and make repairs, if necessary, to ensure that the hearer is able to anticipate the appropriate familiarity level of the upcoming referent.

Epstein (2002) proposed a broader discourse-based framework to provide a unified account of all uses of the definite article. After rejecting the guiding assumption of most previous studies of the definite article in English which posited that definiteness is a matter of identifying and distinguishing individual entities, Epstein argued that "*the* triggers the establishment of connections between various sorts of cognitive domains and the mental entities within those domains" (p. 334). In other words, the definite article creates "accessibility" of a discourse referent so that the hearer can cognitively access the referent. In this framework, the definite article is considered a grammatical signal that triggers construction and retrieval of mental entities in the hearer's mind.

To support his theory, Epstein examined naturally occurring data collected from various sources ranging from classical literature to talk-show transcripts. The analysis of such data yielded four discourse functions of the definite article.

(1) To pick out, or distinguish, a discourse referent that the speaker assumes is uniquely

identifiable to the hearer: because a uniquely identifiable or familiar discourse

referent is, by the same token, accessible, the various sources of definiteness give rise

to access paths in a straightforward manner,

(2) The discourse prominence of an entity: a common literary strategy of employing a

definite description to introduce an important entity at the start of a narrative,

(3) The entity's status as a role function: NPs that designate roles are used to refer to a

fixed property, not to a particular individual (e.g. *the president*),

(4) When an entity is presented from a third person's point of view: NPs with *the* indicate

that a discourse referent is accessible from the point of view of a third person, either a

fictional narrator or a discourse protagonist.

(e.g. the opening sentence of Hemingway's *A Farewell to Arms*: In the late summer of

that year we lived in a house in a village that looked across *the river* and *the plain* to

the mountains.)

Epstein held that the type of access path is generally under-specified by the grammar,

and it is the responsibility of the hearer to single out one interpretation out of several possible

interpretations in a given context. Epstein's (2002) approach is cognitive and discourse-based,

one that assumes active engagement of both the hearer and the speaker in the construction and

interpretation of the definiteness status of a referent.

Research studies that focus on how articles are employed in the construction of a

discourse point to how speakers of English are aware of discourse functions of articles. Speakers

manage referential coherence by selecting articles based on various discourse-related factors to

achieve successful communication. When L2 learners of English do not fully understand these

discourse-related features of articles, their knowledge of English articles may be limited. As Pica

(1983) and White (2010) reported, the textbook grammars are not effective pedagogically as they

do not reflect actual usage of English speakers. For instance, Pica's (1983) data revealed that one

of the most extensively taught rules for article use in English – namely *a* for first mention and *the* for second mention – is not always followed by native speakers of English but is heavily influenced by other discourse factors. These findings suggest the importance of discourse factors to the teaching and learning of English article use.

   2.2.5. History of the Zero Article.

   Most empirical studies on L2 acquisition, learning, and teaching of English articles are based on the underlying assumption that the English article system consists of three members – the definite article *the*, the indefinite article *a/an* (henceforth *a*) and the zero article. The zero article is a grammatical term to account for all the nominal instances in which neither the definite article nor the indefinite article is used. In the previous sections of this chapter, discussions mainly centered on the definite article, and the indefinite article was dealt with in a supplementary manner. Up to this point, however, the zero article did not receive much attention. Unlike the other members of the system (*the* and *a*), the zero article is assumed to lack an overt morphological representation. In this sense, the zero article has taken a history different from the two overt articles. This section provides a brief review of how the notion of the zero article has emerged and whether it is a valid theoretical and linguistic construct that reflects the reality of English language use.

   The existence of the zero article began to be posited around mid-twentieth century. Until then, English was considered to have two articles – the definite article and the indefinite article. For instance, William Cobbett (1906) reflected the traditional grammar approach to the use of English articles in the 19th century. Cobbett attempted to teach his young son "those principles and rules according to which articles are varied in order to make them suit the different

circumstances which they are used to express" (p. 27). Here, he focused on the usage of the two articles. First of all, *a* and *an* "do not *define*, or *determine*, what particular object is spoken of" (p. 27). Cobbett held that indefinite articles only point out the kind of person or thing without defining what. Cobbett also described that the indefinite article can be used before the singular number noun only, such as *an elephant* and *a vowel*.

Next, the definite article *the* "determines the particular object of which we speak" (Cobbett, 1906, p. 28). According to Cobbett, there are other uses of the definite article in addition to its role of defining an object. The definite article is sometimes used "before words expressive of degrees of comparison" as in superlative phrases such as *the best*. Also, Cobbett explained that people use the definite article when they use a singular number noun to refer to a whole class. Cobbett's explanations of the grammar rules for articles included only those rules for the definite and indefinite articles without reference to instances of the zero article.

Similarly, Christophersen (1939), whose work was reviewed in detail in Section 2.3.3, did not invoke the notion of the zero article. In his book, Christophersen recognized the existence of only two articles, and made the distinction between any article (either *the* or *a*) and no article by proposing the actualization theory of articles. When an article is added, "a substantive, from being the name of a mere idea, is turned into the name of something actual and real" (p. 54). For instance, *house* is only a concept without an article, whereas *a house* or *the house* refers to a real, actual house. Christophersen observed that any article seems to change the meaning of the noun in a certain sense, by "actualizing" the basic concept in question so that it is no longer merely "in the abstract". Both *the* and *a* actualize a noun by setting conceptual limits to it: "to receive an article, a word must stand for something viewed as having precise limits" (p. 69). Thus, *cake*, without any article, is opposed to both *a cake* and *the cake*, in that the latter two refer to

particular limited instances of the general concept 'cake'. In this sense, the true opposition in Christophersen's theory is between any article and no article.

Then, in mid-twentieth century, the notion of the zero article began to take shape. In this regard, Berezowski (2009) explained that the first significant work in which the zero article was used to describe English grammar was *A Modern English Grammar on Historical Principles* by Jespersen (1949). It was, in fact, not Jespersen who decided to use the term 'zero article' but his student Niels Haislund who completed Jespersen's work after his death in 1943. At that time, Haislund was under the influence of Jakobson who posited that in contexts showing a systematic opposition between an overtly expressed element of grammar and its absence, the latter should be recognized as a real member of the pair and called the 'zero element' due to its lack of overt expression. This structural linguistic approach was quickly and conveniently applied to well-known data previously described in terms of absence of articles.

In discussing the meaning of the zero article, it is necessary to revisit Jespersen's (1949) stages of familiarity. As discussed in Section 2.3.3, the first stage is the stage of complete unfamiliarity, and either the zero article or the indefinite article is used for this stage. In the second stage, the familiarity of a referent is nearly complete, and the definite article is used for this. Finally, in the third stage, familiarity of a referent is so complete that no article is needed. Thus, it is posited that the zero article is used when a referent is completely familiar. In this schema, one may wonder how the zero article, posited as a single grammatical entity, can be used in the opposite instances of complete unfamiliarity and complete familiarity. As such, Berezowski (2009) argued that "the zero article inspires little scholarly interest because it is presumed to be both definite and indefinite at the same time, which makes it an internally contradictory entity that can be hardly theorized about with any success rate or put to any other

viable uses" (p. 2). In other words, most scholars accepted the existence of the zero article as a legitimate grammatical entity and moved on with other research endeavors without attempting to account for the validity of the zero article.

One solution that was proposed to deal with this contradiction was to posit two covert articles. The rationale for positing two covert articles is quite simple. Since it was impossible to posit a single zero article to account for a range of data that was too wide to be adequately captured in one formula, it was only logical to divide various usages into two parts and try to treat them separately. For instance, Yotsukura (1970) differentiated between the zero article used for indefinite mass and plural nouns, and the 'no-article' form, as used before singular proper nouns. Yotsukura (1970) presented the following expressions to argue for the distinction between 'zero-article' and 'null-article'.

I like (*zero Art*) cheese. vs.   I like the cheese.
(cheese in general)   (a particular kind of cheese that
                                                             is presented there or talked about)
I like (*null Art*) London. vs.   *I like the London.

In the above example, the absence of the article before *cheese* is considered 'zero article' which usually occurs before mass nouns and abstract nouns. The case of *London* is considered a 'null article' which occurs with proper nouns. However, one can easily think of a sentence where an article is used before a singular proper noun, as in *I like the London of the 60's*. Such counter examples show that the distinction between 'zero' and 'null' articles needs further analysis with actual language data.

Similar to Yotsukura's approach, Chesterman (1991) proposed that the English article system consisted of five members – zero, *some*, *a*, *the*, and null – and, in doing so, argued for two covert articles: the zero article and the null article. As discussed earlier in Section 2.2.3,

Chesterman explained the semantics of articles in terms of locatability, inclusiveness, and extensivity. The five members of the article system can be described in terms of these key properties.

Table 2. 1. Key article properties according to Chesterman (1991, p. 68)

Article	Locatable	Inclusive	Limited extensivity
zero	-	±	-
some	±	-	+
a	±	-	+
the	+	+	+
null	+	+	-

According to Chesterman, the zero article is compatible with any nominal whose referent is a non-locatable set that has at least two members, and the null article is compatible with any nominal whose referent is a locatable one member set. For instance, the zero article is used in *There are ∅ security guards outside his mansion* since the referent (*security guards*) has more than one member of the set. On the contrary, the null article is considered to be used in *∅ John was my first love* because *John* is a proper noun whose referent is a one member set. What the covert articles have in common is that they are both used for nominals with unlimited extensivity, which assumes that their referents are abstract categories. However, the locatability of individual members of the set depends on their number. If the set contains only one member, then the locatability of its member is guaranteed, but if the set contains more than one member, such identification of the referent is not guaranteed. This difference in locatability between the null and zero articles is indicated in the above table. However, the most significant weakness of Chesterman's approach is that a number of English singular proper names require the use of the

definite article, as in *the Indian Ocean*.

Criticizing the unchallenged assumption of the existence of the zero article, Berezowski (2009) argued that "the zero article is an artifact of the structural paradigm of language research that has outlived the heyday of the theory that spawned it and remains in use merely as a fig leaf concealing the fact that a number of grammatical issues have failed so far to attract the interest of scholars working in later and more adequate frameworks" (p. 2). In accounting for instances in which no overt article is admissible, Berezowski proposed the incomplete grammaticalization model. According to this model, articles in English have been undergoing stages of grammaticalization. In other words, articles have been used in more extensive grammatical contexts and environments over time, and the expansion of grammatical environments for article use has been occurring along a predefined path. The evolution of English articles is still incomplete.

Therefore, Berezowski explained that "the gaps in the use of the English definite and indefinite articles are a direct consequence of incomplete article grammaticalization" (p. 48). Grammaticalization proceeds along a predefined path, so at the macro level of the history of language, the process is predictable. However, particular grammaticalization steps are not always orderly and consistent. Therefore, along the process of grammaticalization, gaps may occur in which overt articles have not yet expanded to certain sets of grammatical environments.

Further, Berezowski explained that only 8% of languages have two-article systems. In languages that do not formally mark definiteness, the identification of definite and indefinite nominals is a pragmatic issue that requires some effort on the part of the hearer. However, the absence of articles "does not impair the efficiency or precision of communication" (p. 50). For instance, Kim (1990) pointed out that Korean does not have articles indicating definiteness or

indefiniteness. He stated that "although definiteness is indicated by demonstratives in some cases, the distinction between definite and indefinite, in general, is not made in Korean" (p. 895). As such, interpretation of definite and indefinite reference is often constructed by the hearer through inference and contextual cues in a given discourse.

In the case of English, articles originated from a demonstrative and a numeral, and as the process of grammaticalization began, article usage has been expanded to other non-referential and non-numeral environments. The process is still on-going, and along the way, contexts in which overt articles are not expressed have emerged. In other words, environments in which overt articles have not yet expanded may remain as contexts in which no article is required, and such contexts cannot be characterized in a coherent manner. Based on these observations, Berezowski rejected the notion of the zero article and argued that the English article system has only two members.

This current study has been informed by Berezowski and it assumes that the English article consists of two members: the definite article *the* and the indefinite article *a*. There are several reasons why this particular view of English articles was adopted in the current study. Firstly, Berezowski's perspective incorporates both the diachronic development and the synchronic status of English articles in a more accurate manner. As discussed in the previous segment, the two overt articles (*the* and *a*) originated from semantically and morphologically distinct forms. Over time, the environments in which these articles were used have expanded following separate paths of grammaticalization without reference to any 'covert' article. Synchronically, the two overt articles denote distinct semantic and pragmatic attributes of nominals in English. As discussed in Sections 2.2.3 and 2.2.4, *the* is concerned with the definiteness of nominal reference perceived by the speaker and the hearer while *a* is an

indication of singularity and unity of a nominal. Considering the historical and current characteristics of articles in English, the present study recognizes only two articles in English without the zero article.

Secondly, the two-article model offered analytical advantages in analyzing and interpreting instances in which no overt article was produced. On the surface, it is impossible to distinguish between instances of the zero article and instances of article omission. In this regard, the challenges in analyzing and interpreting data on the use of the zero article have been described in the literature (Master, 1997; Parrish, 1987; Zegarac, 2004). For instance, Master (1997) argued that researchers cannot tell the difference between the zero article and non-use or omission of article because both instances look identical on the surface level. This analytical concern can be addressed by adopting Berezowski's framework that recognizes only the two overt articles. Thus, in the current study, an instance in which an overt article was not produced as required was analyzed as an instance of article omission. In turn, when an overt article was not used and it was deemed appropriate in a given context, such an instance was analyzed as an accurate instance, without attempting to distinguish between an instance of non-use and an instance of the zero article. Furthermore, previous research (Ekiert, 2010; Park, 2006, Zegarac, 2004) reported that L2 learners whose L1s do not have an article system tend to achieve relatively high accuracy in contexts in which an overt article is not required. The high accuracy in these instances is often attributed to positive L1 effect, as suggested by Zegarac (2004). Recognizing this tendency in L2 acquisition of articles, the current study focused only on the two overt articles (*the* and *a*) as they are the forms that caused most difficulties to Korean learners of English.

In this section, the English article system was reviewed and discussed in detail. Now, we

review what has been documented in previous research with regard to how second language learners understand, acquire, and use English articles.

2.3. Research on Second Language Acquisition of English Articles

Mastery of English articles has been recognized as a significant challenge for second language learners. English articles constitute a complex system where morphological forms are matched up with multiple sets of semantic distinctions and pragmatic functions. For this reason, there has been extensive research on the topic in the field of second language acquisition. In this section, research studies with regard to second language acquisition of English articles are discussed in detail. The first part presents a review of studies that explored and described patterns of article use by L2 learners of English. The second part discusses studies that have attempted to propose reasons for difficulties associated with L2 use and acquisition of English articles. Finally, research on the teaching of articles to L2 English learners is discussed.

2.3.1. Descriptive Studies.

Some early studies on second language acquisition of articles used the concept of the Suppliance in Obligatory Contexts (SOC) (Brown, 1973), which was widely used to determine if a particular linguistic feature is acquired. These studies looked at whether learners of English supplied articles in contexts that required articles and determined if the learners reached a level of acquisition of articles. However, this approach did not take into consideration various semantic and discourse characteristics of different articles in their analyses and thus failed to provide an accurate account of how L2 learners of English learn and use different articles in English.

Scholars began to conduct second language research dedicated to the use and acquisition of English articles in the 1980s. In particular, Huebner (1979) made an important contribution by proposing a method of analysis called "dynamic paradigm" (p. 22). Huebner collected an oral corpus produced over a period of 12 months by a 23-year-old male Hmong refugee from Laos, living in Honolulu and learning English in a natural setting without formal instruction. Based on the assumption of interlanguage that "language is systematic but dynamic, and that variation is the precursor to change" (p. 22), Huebner designed the dynamic paradigm of English articles to examine variation in interlanguage. This paradigm accounts for article use in all pre-noun positions in terms of Specific Referent (SR) and Assumed Known to Hearer (HK).

1. +SR, +HK (*the*)
   a. Unique reference
   b. Conventionally assumed unique reference (e.g. *The President*)
   c. Referent physically present
   d. Referents previously mentioned in discourse
   e. Specific reference assumed known to hearer
2. –SR, +HK: Generics (e.g. *cancer*, *peace*) (*a*, *the*, zero)
3. –SR, –HK (*a* / zero)
   a. NPs in scope of negation
   b. NPs in modal, irrealis scope
   c. NPs in scope of questions
   d. NPs with indefinite number
4. +SR, –HK: First mention of NP [+SR] in a discourse (*a* / zero)

According to Huebner (1979), in Standard English, noun phrases in category 1 are marked with *the*, while those falling in categories 3 and 4 are marked with *a* or zero. Category 2 represents generics, which can be marked with *a*, *the*, or zero. By applying this model of analysis to the oral corpus data from the Hmong learner of English, Huebner showed that the participant's hypotheses about the target language feature was under continuous revision.

Huebner's dynamic paradigm has had a significant influence on subsequent studies that described patterns of article use by L2 speakers of English (Parrish, 1987; Tarone & Parrish,

1988; Thomas, 1989; Young, 1996, among others). His diagram provided a simple but comprehensive frame of analysis that made it possible for researchers to consider both the aspect of specificity of a referent and the aspect of the hearer's understanding of a referent in their analysis of article usage.

One of the studies that utilized Huebner's (1979) semantic paradigm was carried out by Tarone and Parrish (1988), a study that conducted detailed quantitative and qualitative analysis of data collected by Tarone in 1985. The goal of the study was to investigate how different task types influence English article production by L2 learners of English in their interlanguage. In Tarone's (1985) study, ten native speakers of Japanese and ten native speakers of Arabic performed three tasks in English. Both languages do not have article systems. The tasks administered in this study are the following:

(1) a written grammaticality judgment task where the participants were asked to identify ungrammatical sentences and rewrite them correctly
(2) an oral interview with a native speaker of English, focusing on the participants' field of study and plans for the future
(3) an oral narration task where the participants were asked to look at a sequence of events on a video screen and narrate a story to a non-native listener

In order to investigate relationship between task types and semantic types of noun phrases, Tarone and Parrish used Huebner's (1979) semantic types of NPs based on [Specific Referent (SR)] and [Hearer Knowledge (HK)]. Tarone and Parrish developed the following three null hypotheses and conducted a *t*-test.

(1) There will be no difference in the proportion of NP Types produced by learners on the two oral tasks.
(2) There will be no overall difference in learners' accuracy rates in article production related to NP Type.
(3) There will be no difference in accuracy rates in article production related to NP Type, when learners move from one oral task to another.

The results of the statistical analysis rejected the first hypothesis. In the interview task, [-

SR, +HK] contexts, [+SR, +HK] contexts, and [+SR, -HK] contexts were produced with similar frequencies. On the other hand, in the narrative task was Type 2, a majority of NP contexts produced were [+SR, +HK] contexts, and [-SR, +HK] contexts were almost non-existent. As for the second hypothesis, the accuracy in the use of articles in [+SR, -HK] contexts was significantly lower than in the [+SR, +HK] contexts. Finally, the third hypothesis was barely upheld. An interesting finding was that the accuracy of article use in [+SR, +HK] contexts was greater in the narrative task than in the interview task. This finding indicated that not only [+SR, +HK] contexts, which are specific referents known to the hearer, were used far more frequently in the narrative task, but also the accuracy with which articles were produced in these contexts was higher in the narrative task than the interview task. In other words, the participants were able to employ the definite article more correctly in [+SR, +HK] contexts when they narrated stories.

Another interesting finding from Tarone and Parrish's (1988) study is that articles were produced with less accuracy in the grammaticality judgment task and far more accurately in the oral narration task. The scholars explained that "effective story-telling requires that the narrator keeps track over time of persons and objects important to the story-line" (p. 34). One use of [+SR, +HK] contexts is for referents previously mentioned in the discourse. The increased accuracy of article use in [+SR, +HK] contexts in the narrative task can be attributed to the communicative pressure on the narrator to maintain referential coherence. In conclusion, their study demonstrated how different task types may influence the use of articles by ESL learners and illustrating that discourse functions may influence interlanguage variation with regard to article usage.

Several researchers made modifications to Huebner's semantic distinction of English articles. For instance, Parrish (1987) pointed out a need for modifying Huebner's (1979)

39

paradigm and made changes to the analysis frame. In analyzing oral production data of a 19-year-old Japanese woman collected for a period of four months, Parrish modified Huebner's (1979) system by adding proper nouns, idiomatic expressions, quantifiers, and demonstratives, which were omitted in Huebner's study (1979). Parrish's study found that the definite article *the* was being acquired earlier than *a*, and that the rate of accuracy of *a* decreased when the use of *one* increased. The study suggested that the participant adopted a strategy of using a quantifier or overusing the definite article for nouns that would otherwise require the indefinite article. Initially, the subject of Parrish's study correctly used idiomatic expressions such as *all day* or *at night*, but about two months into data collection, she began to question these structures and started to insert the definite article, as in *all the day* and *at the night*. Parrish (1987) argued that this tendency supports Hakuta's (1976, cited in Parrish, 1987) notion of internal consistency, in which a learner tries to keep related linguistic forms within his/her system consistent with one another.

Young (1996) adopted Huebner's classification of noun contexts by the referent semantics represented by Specific Referent (SR) and the discourse features of contexts represented by Hearer Knowledge (HK). However, unlike Huebner who examined only the definite article, the indefinite article and the zero article, Young (1996) proposed that the article system in the interlanguage of L2 learners consisted of five members: definite article, indefinite article, zero article, demonstratives, and pre-nominal modifiers (e.g. *some*, *any*, *no*). More specifically, the researcher attempted to study the influence of interlanguage development, noun countability, noun reference, discourse marking, and L1 transfer on variation in articles in English interlanguage. The study included three Czech speakers and three Slovak speakers at different stages of English language learning.

The study revealed important findings concerning how these learners would map L2 article forms to their functions. First of all, accuracy in article use increased roughly in proportion to accuracy in other aspects of English as measured by TOEFL scores. Second, although the participants rarely used indefinite articles, when they did so, they used them to mark [-HK] nouns. As for definite articles, learners at the low proficiency level showed a tendency for over-generalization regardless of discourse contexts. As proficiency level increased, learners began to use definite articles to mark unique and specific reference. Third, with regard to noun countability and number, the participants systematically used indefinite articles to distinguish number and countability. Young (1996) explained that a mapping between L2 form and function in this context seemed to be easy for these learners, since number is clearly marked in their L1. Finally, the study found that demonstratives such as *this* and *that* were used widely to indicate anaphoric reference. Since demonstratives are widely used in Czech and Slovak to mark definiteness of a noun referent, this tendency was understood as an effect of L1 transfer.

Merits of Huebner's diagram (1979) have been attested by its extensive use in the above studies. While the diagram incorporates complicated semantic and discourse aspects of English articles, it is still easily applicable to empirical data thanks to its neat and easily comprehendible representation. For these reasons, Huebner's diagram has been a framework for researchers to document interlanguage development of second language learners. However, despite its usefulness for descriptive analysis, Huebner's framework does not provide satisfying explanations for observed variations in the use of articles by various L2 learners. A major issue with his framework is that the Huebner model is a classification system based on semantic features of nominal phrases rather than individual articles. As such, the model does not adequately depict form-meaning correspondences of the article forms.

In addition to Huebner's diagram, other models of article classification have been used by scholars to describe article use by L2 learners of English. For example, Liu and Gleason (2002) took Hawkins' (1978) classification of non-generic types of the definite article and divided them into four categories of non-generic use of the definite article. The generic use of the definite article is when the definite article is used for generic referents, as in <u>*The whales* are the</u> *largest mammals on earth*. All the other uses of the definite article were considered by Hawkins as non-generic types and the non-generic types were classified into the following categories:

(1) Cultural use: when a noun is a unique and well-known referent in a speech community
    a. *larger situation use relying on specific knowledge*
    b. *larger situation use relying on general knowledge*

(2) Situation use: when the referent of a first-mention noun can be directly or indirectly sensed in the immediate situational context
    a. *visual situation use* when the speaker and the hearer can see the referent
    b. *immediate situation use* which is similar to type 2 but the referent may not be visible

(3) Structural use: when *the* is used along with certain types of modifiers
    a. *unfamiliar use in NPs with explanatory modifiers* when a first-mention noun has an identifying modifier
    b. *unfamiliar use in NPs with non-explanatory modifiers* when adjectives such as *same* and *only* are used

(4) Textual use: when a noun is subsequently mentioned or related to a previously mentioned noun
    a. *anaphoric use* when something is mentioned subsequently
    b. *associative anaphoric use* when a first-mention referent is related to a previously mentioned noun

Using the above classification, Liu and Gleason (2002) explored whether different uses of the definite article in the non-generic contexts present different levels of difficulty for ESL learners and whether these uses are acquired in the same order. In the study, 41 low-, 49 intermediate- and 38 advanced-level ESL students were asked to read 91 sentences and insert *the*

wherever they deemed necessary. The 91 sentences contained 60 obligatory uses of the definite article and 40 zero articles for control items. Statistical analyses of the participants' performance yielded two major findings. First, the four non-generic uses of the definite article posed a varying degree of difficulty to the learners, and they were not acquired together. The data suggested that ESL learners seemed to acquire situation use first, the structural and textual uses in the middle, and the cultural use last. Secondly, the participants' omission of *the* in obligatory contexts decreased significantly as their English proficiency improved. However, the overuse of *the* presented a different picture. ESL learners were found to over-generalize the definite article as their proficiency improved from beginner to intermediate level. The tendency to overuse *the* began to diminish with the improvement of English proficiency as they became more aware of appropriate uses of the definite article.

In a similar study, Kim and Ahn (2003) examined the acquisition of the definite article by Korean high school students. They used the model of non-generic use of the definite article proposed by Liu and Gleason (2002), which is based on Hawkins' (1978) classification, to classify the various uses of non-generic *the*. The instrument used in the study consisted of 87 sentences containing 58 deleted obligatory uses of *the* (13 cultural, 13 textual, 16 situational, and 16 structural) and 36 null articles. The participants included 40 low-, 40 intermediate- and 33 advanced-level high school students. They were instructed to read the sentences and insert the definite article *the* wherever they believed necessary. The findings of the study showed that the various uses of *the* presented different levels of difficulty for Korean high school students and that they seem to acquire textual use first, situational use second, cultural use third, and structural use last. The order of acquisition shown in the study was different from what was found in the study by Liu and Gleason, which found that the cultural use was most difficult. However, Kim

and Ahn (2003) did not attempt to explain why the Korean high school students seem to go through a different order of acquisition.

Kim and Ahn's (2003) study duplicated Liu and Gleason's (2002) research design with a different population (i.e., Korean learners of English), and the order of acquisition was different from the findings of Liu and Gleason's study. This concern raises a question of whether the four categories of non-generic use of the definite article are a valid way of classifying various uses of the definite article. Another limitation of the study is a lack of explanation as to why certain uses of *the* are deemed more difficult to Korean high school students. It may be related to the fact that these students are situated in an English as a Foreign Language (EFL) context where exposure to authentic language use is extremely limited.

It has been extensively documented that learners of English whose L1s do not utilize articles tend to use articles less accurately than those from L1s with similar article systems (e.g., Ekiert, 2010; Park, 2006; Robertson, 2000; Young, 1996). The following segment reviews several descriptive studies that focused on how L2 learners of English from such L1 backgrounds employ articles. These studies mainly deal with possible L1 effects on the accurate production of articles by these learners whose L1s do not have a similar article system, such as Chinese, Japanese, Korean, and so on.

In a study on the topic of learning English articles by Korean speakers, Lee (1997) carried out a quantitative analysis of the frequency of English article deletion in compositions by Korean learners of English as a Foreign Language (EFL). In the study, 49 university freshmen in Korea were asked to write a spontaneous essay on the topic of 'My Freshman Lab Class' during a 50-minute time frame. With social factors in consideration, the author tested four hypotheses:

(1) the rate of English article deletion will be considerably higher than the rate of article

suppliance
(2) males will delete the English article more than females
(3) subjects with rural backgrounds will delete the English article than those with urban backgrounds
(4) the deletion rate of the English article will be higher if the articles precede a 'modifier + noun' rather than directly preceding a noun

The analysis of the writings rejected the first hypothesis with only 129 tokens out of a total of 638 to be instances of article deletion. Lee (1997) suggested that the low deletion rate might be attributed to the fact that it was a writing task that allowed learners time to correct their mistakes. Also, the gender factor and the factor of social status (i.e., learners with rural backgrounds as opposed to urban backgrounds) proved statistically insignificant. That is, article use was not found to be influenced by gender or social economic background. The fourth hypothesis, however, was supported by the data in that it suggested a slight tendency to delete the article more frequently when it preceded an adjective or an adverb. Lee (1997) suggested that this tendency could be explained by L1 transfer since the role of the article is often expressed by modifiers in the subjects' L1.

One major issue with Lee's study is the concept of 'deletion', which implies the existence of something that can be deleted. However, in the case of Korean learners of English, they are likely to omit articles rather than delete them. That is, given that Korean does not have articles, instances of where no articles were used should be interpreted as omission errors. In addition, Lee did not sufficiently explain the rationale behind the inclusion of gender and social background as potential variables that may influence article choices, considering that there has not been any research that indicated possible social and gender effects on the use of articles.

Based on the assumption that second language learners whose first languages contain an article system greatly differ from those whose L1s do not contain such a system, such as

45

Japanese, Mizuno (1999) attempted to uncover the process of English article acquisition by Japanese learners of English. The researcher first summarized his previous approach to interlanguage analysis put forward in 1988 and examined new data from experiments conducted with approximately 2,000 Japanese EFL students at seven universities in Japan.

In the previous study, Mizuno (1988) established that there are at least five types of errors in the use of articles that Japanese learners are likely to produce (p. 3):

(1) errors in co-occurrence: juxtaposition of articles and other determiners
(2) word-order: inversion of the order of articles and succeeding adjectives
(3) under-extension: omission of articles
(4) over-extension: the use of article instead of *zero*
(5) substitution: *a* used instead of *the* or vice versa

The resulting data from a sample of 800 high school and university students in Japan supported Mizuno's (1988) claim that the first two types of errors – errors in co-occurrence and word-order errors – would occur at the beginning level and decrease significantly at early intermediate level since they are rather mechanical errors. The other three errors persisted through the proficiency levels.

The results of the analysis of newly collected data from 2,000 Japanese EFL students informed the researcher about the following constraints affecting Japanese adult learners' performance with English articles:

(1) confusion between the concept of zero and null
(2) ignorance of the concept of the zero article in English
(3) a lack of awareness of the boundary between the discrete and continuous attributes of nouns
(4) limited opportunity for practice and exposure to meaningful expressions in English and the authentic rhythms of English in everyday life
(5) a complete lack of awareness of deictic factors which constrain the article used.

Based on the above findings, Mizuno (1999) discussed relevant pedagogical implications for speakers of languages that do not have an article system.

In another descriptive study of article use by Korean learners of English, Park (2006) recruited 33 Korean adult learners of English who were students in the U.S. at the time of the data collection. They were divided into three proficiency groups. The participants carried out a written fill-in-the-article task and produced a written essay on the topic of things they liked about living in the U.S. The purpose of this study was to investigate if levels of English language proficiency could predict Korean learners' accuracy in the use of English articles.

After conducting quantitative analyses of the written production of articles by the participants, Park (2006) concluded that overuse of the zero article was the most frequent and distinct characteristics of Korean learners' use of English articles. Moreover, he argued that levels of English proficiency did not seem to affect the Korean students' overuse of the zero article. In other words, the frequencies with which participants overused the zero article did not vary significantly among the three proficiency groups. Another interesting finding in Park (2006) is that, unlike what has been documented by scholars such as White (2003), Parrish (1987) and Zegarac (2004), the overuse of *the* in both obligatory and nonobligatory contexts was not frequent in this study. In short, among Korean learners of English, omission of *the* was more evident than over-use of the form in their written production.

A major limitation of Park's study is that Park did not distinguish article omission from over-use of the zero article. On the surface, these two behaviors may look identical, but they may originate from different motivations. In other words, article omissions may be indicative of L1 transfer effect since Korean does not have a similar article system. In turn, over-use of the zero article is suggestive of an active choice of the zero article based on inaccurate judgment of a particular nominal context. The challenges in analyzing and interpreting data on the use of zero have been attested in the literature (Master, 1997; Parrish, 1987; Zegarac, 2004). For instance,

Master (1997) argued that researchers cannot tell the difference between the zero article and non-use or omission of article because both instances look identical on the surface level. In this regard, Park (2006) did not adequately discuss the risk of mixing the two cases in interpreting instances in which no article was produced by the participants.

While the above studies analyzed patterns of article production, Butler (2002) examined L2 learners' metalinguistic knowledge with regard to the use of English articles by collecting data from Japanese learners of English. The research questions were three-fold:

(1) types of metalinguistic knowledge L2 learners employ in selecting English articles
(2) the manner in which learners understand the features of the English article system at the metalinguistic level
(3) relationship between the use of metalinguistic knowledge and the level of proficiency in English

A total of 100 participants were recruited for the study, among which 80 of them were Japanese students (20 each in three proficiency groups living in Japan and the remaining 20 students studying in a U.S. university), and the remaining 20 were American college students as a control group. The participants took a fill-in-the-article test that contained 100 items (20 each for Huebner's four-way classification and the remaining 20 for idiomatic and conventional uses). Immediately after the test, the Japanese participants were asked to provide the author with reasons for their article choice on each item.

The participants' performance on the test showed that learners with higher proficiency exhibited more target-like article choices, though there remained a large gap between the native English speakers and the most advanced Japanese learners. With regard to metalinguistic knowledge of the English article system obtained from the study, the Japanese learners' hypotheses about articles were classified into three categories. The context-insensitive hypotheses were found to be heavily influenced by textbooks and teachers and include such rules

as "when an object or event is specific, *the* is used" or "when a noun is countable, *a* is used". Learners might apply these rules without having a clear understanding of how hearer knowledge, countability and other contextual factors are related to the English article system. The second type of learner hypotheses about articles were formulated and used by Japanese learners of English when they showed sensitivity to inappropriate contextual cues. The final type of hypotheses was exercised by learners who were sensitive to a range of relevant contexts.

With regard to proficiency levels, the first type of hypotheses was found to be used more extensively among low proficiency learners as they were influenced by textbooks, teachers and so on. As the proficiency improves, learners seem to realize that context-insensitive hypotheses failed to provide them with the correct article choices and they start to develop ad hoc rules by linking article usage to inappropriate contextual cues. As learners developed hypotheses to capture the actual uses of articles as much as possible, they tended to create certain collocation rules. For instance, one participant observed many instances where no article was used after the preposition *of*. Then, he developed a working hypothesis that no article is allowed after *of* and started to apply it elsewhere regardless of other contextual variables. Finally, advanced learners seem to become more sensitive to relevant contextual cues when correctly using articles. However, Butler (2002) found that such hypothesis patterns were not always clear-cut as learners with different levels of proficiency seem to mix various types of hypotheses about article use.

The significance of Butler's (2002) study lies in the use of qualitative interviews along with article use test results to examine metalinguistic awareness of Japanese learners of English. The study reported that Japanese learners initially rely on the rules they learned from textbooks and instruction with regard to article choices as they do not have any equivalent system in their native language. However, these rules often fail them, after which learners start to form and test

various hypotheses. As informed greatly by Butler (2002), this present study also investigated

metalinguistic hypotheses of Korean learners of English at different proficiency levels.

2.3.2. Studies that Explain Difficulties in L2 Acquisition of English Articles.

Thanks to the descriptive and exploratory studies conducted in the 1980s and 1990s, a

wealth of knowledge about patterns of article use and acquisition by L2 learners of English has

been accumulated. Since then, attempts have been made to provide explanations for the patterns

of article use and acquisition that have been reported in descriptive studies. To this end, a variety

of theoretical constructs and hypotheses have been proposed to account for challenges associated

with L2 acquisition of articles by learners whose L1s do not have a formal marking of nominal

definiteness.

In an attempt to provide an explanation for a tendency found in Chinese speakers to omit

articles in obligatory contexts, Robertson (2000) argued that difficulties experienced by Chinese

learners of English are associated with the challenges of transitioning from a discourse-oriented

language to a syntax-oriented language. To test his hypothesis, Robertson conducted a study on

18 speakers of Chinese origin, who were postgraduate students at a major university in the UK

with an advanced level of English proficiency. The participants took part in a collaborative

problem-solving task. For data analysis, Robertson used Hawkins' (1978) model that categorizes

definite and indefinite NP contexts. The summary of Hawkins' model is as follows.

Definite NP environments (examples are from Robertson, 2000)
(1) Anaphoric use of referring NP: subsequent mention of the first-mention NP.
      A: And then after that you draw a square with the red…
      B: Square?
      A: Yeah, a square.
      B: What does the square draw like?
(2) Immediate situation use of referring NP: the object referred to is present in the

immediate situation.

> A: *Am I to use the blue, blue…?*
> B: *No, no, no, <u>the red pen</u>.*

(3) Larger situation use of referring NP: the NP referred to can be uniquely identified on the basis of knowledge which is shared by the speaker and hearer.

> A: *Left and upper, upper half of the, of the, er, of the paper spread wide. Okay, towards <u>the center</u>, top center…*
> B: *Top center, top center*.

(4) Head noun of an associative clause NP: dependent on knowledge of a conventional association between the two NPs, e.g. <u>the front</u> of the house.

(5) 'Unexplanatory' use of definite NP: e.g. *the same N*, *the first N*, *the best N*, etc.

(6) NP with nominal modifier: e.g. *the number 3*, *the letter A*.

(7) NP with establishing relative clause: the definiteness of the NP is established through the use of a post-modifying relative clause.

<u>Indefinite NP environments</u>

(1) Use of NP in existential prediction: typically introduced by the phrase *there is/are*, or through the use of the verbs *have* or *got*

> A: *All right, at the top right hand corner of the paper <u>there is a number three</u>.*

(2) Use of NP as object of transitive verb or complement of copulative construction: first mention of noun

> A: *Then you draw <u>a horizontal line</u> using your blue pen.*

(3) Generic use of singular NP

> A: *What do you mean by the square?*
> B: *Square, like <u>a floor</u>, you know?*

The coding of 1,884 noun phrases based on Hawkins' taxonomy showed an overall rate of 78% suppliance of articles in obligatory contexts. With regard to the remaining 22% of contexts where the participants omitted articles, Robertson (2000) claimed that many of these instances of non-suppliance can be attributed to the difficulty associated with moving from a grammar of Chinese which is "discourse-oriented" to a grammar of English which is 'syntax-oriented" (p. 169). In a discourse-oriented language like Chinese, certain grammatical features such as definiteness, number and tense, are not marked overtly when the relevant information can be recovered from the context. In contrast, in a syntax-oriented language like English, these grammatical features must receive overt morphosyntactic realization. Under this framework,

three general principles in linguistics were utilized by Robertson to account for instances where the Chinese participants did not supply appropriate articles: a "determiner drop" principle, a "recoverability" principle, and a "lexical transfer principle" (p. 169).

First, the syntactic principle of 'determiner drop' is analogous to 'pro-drop', an instance in which an NP with definite or indefinite reference need not be overtly marked for definiteness if it is included in the scope of the determiner of a preceding NP. Second, the pragmatic principle of recoverability assumes that in some discourse-oriented languages as Chinese, an NP need not be marked for definiteness if the information encoded in this feature is recoverable from the context. Finally, Robertson identified a lexical transfer principle based on the observation that some learners tended to use demonstratives and the numeral *one* as markers of definiteness and indefiniteness, respectively, just as Chinese does.

Ionin, Ko and Wexler (2004) approached L2 article acquisition from the perspective of semantic parameters under the framework of Universal Grammar, which determine the possible interpretation of articles in languages that have a two-article contrast. The Article Choice Parameter (p. 12) is defined as follows:

> "A language that has two articles distinguishes them as follows:
> The Definiteness Setting: articles are distinguished on the basis of definiteness.
> The Specificity Setting: articles are distinguished on the basis of specificity."

According to Ionin et al., the feature [+definite] reflects the state of knowledge of both speaker and hearer, whereas the feature [+specific] reflects the state of knowledge of the speaker only. Thus, an NP is [+definite] if the speaker and the hearer presuppose the existence of a unique individual in the set denoted by the NP. Meanwhile, an NP is [+specific] when "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" (p. 5). The Article Choice Parameter predicts two possible

patterns of article choice in "two-article languages" (p. 12) cross-linguistically: article grouping by definiteness, as in English, and article grouping by specificity, as in Samoan.

       After analyzing elicitation and production data produced by 30 Russian speakers and 40 Korean speakers whose native languages do not have articles, Ionin et al. concluded that L2 learners have access to both settings of the Article Choice parameter, which may explain why the participants would fluctuate between the definite and indefinite articles. The Article Choice parameter is not relevant to L1s that do not have an article system such as Russian and Korean. As these participants begin to learn English, the parameter needs to be set newly for English. In doing so, they experience a stage where fluctuations occur between four possible settings within the Article Choice parameter before they would ultimately settle for the proper setting of English. In particular, the data analysis showed that a majority of the participants used the definite article for indefinite contexts with specific references, which provided evidence for their article choice based on referent specificity. For instance, in *Peter intends to marry a merchant banker, even though he doesn't get on at all with her*, 'merchant banker' is in an indefinite, specific context because the speaker intends to refer to a particular individual with a noteworthy property. Yet, learners chose the definite article in this type of context. Based on these findings, Ionin et al. (2004) formulated the following hypothesis (p. 16):

> "The Fluctuation Hypothesis:
> a. L2 learners have full access to UG principles and parameter settings.
> b. L2 learners fluctuate between different parameter-settings until the input leads them to set the parameter to the appropriate value."

This hypothesis attempts to explain variation in the article usage by L2 learners in terms of two parameter settings. Learners' overuse of *the* in [+specific, -definite] contexts and overuse of *a* in [-specific, +definite] contexts seem to lend support for the fluctuation hypothesis that L2

learners are affected by the specificity feature in their article choice. As they developed an argument for the fluctuation hypothesis, Ionin et al. excluded the possibility of L1 transfer effect on the ground that both Russian and Korean do not have an article system similar to that of English.

Ionin et al. (2004) introduced two generative hypotheses – Article Choice parameter and the Fluctuation Hypothesis – to account for fluctuations of the use of Articles by L2 learners of English. However, they fail to persuade the reader that the generative framework is better than the interlanguage framework in accounting for the variation that exists in article use. In fact, Ionin et al. acknowledged that even though fluctuation was observed at group level, when examined on the individual level, article choices by a substantial minority (37%) were made on the basis of other unexpected patterns, which their hypothesis could not account for.

Hawkins and his colleagues (2006) noticed limitations of Ionin et al.'s (2004) framework such as those described above, and proposed that an account of variable use of the articles can be given without either appealing to an Article Choice Parameter or to the concept of fluctuation. Hawkins et al. administered to 12 Greek speakers and 12 Japanese speakers a forced-choice elicitation task, which was adapted from the task designed by Ionin et al. (2004). The researchers selected Japanese and Greek for their study because they wanted to test Ionin et al.'s (2004) fluctuation hypothesis in a new population of L2 speakers of English whose L1 does not contain articles such as Japanese, and to test whether speakers of an L1 with articles that mark definiteness such as Greek would show the same tendency to fluctuate. The results reported of the task were very similar to those of Ionin et al. (2004), in that the Japanese speakers tended to use the definite article in indefinite situations with specific referents and the Greek participants did not exhibit such tendency. This finding suggested that fluctuation is not a general L2

development phenomenon. As in the case of Greek learners of English, fluctuation may not occur in L2 learning when a learner's L1 encodes definiteness like English, because the learner's article choice parameter could be already set for definiteness in their L1.

In addition to examining group tendencies between the Japanese and Greek speakers, Hawkins et al. (2006) attempted to explain individual variation observed among the Japanese participants. To this end, they introduced the assumption made in recent L2 studies that "the output of syntactic operations produces strings of terminal nodes that are bundles of features lacking phonological exponents" (Lardiere 2000, cited in Hawkins et al. 2006, p. 19). Because the exponents are inserted only after all syntactic operations have applied, it is called "Late Insertion" (p. 19). In other words, a speaker matches features between a vocabulary item and a terminal node before the insertion occurs. For example, articles are exponents of the category D in English. The terminal node for D will produce bundles of features for native speakers as the following.

[D, +definite, +singular]	[D, -definite, +singular]
[D, +definite, -singular]	[D, -definite, -singular]

However, the contexts of insertion for the phonological exponents are:

a	↔ [D, – definite, + singular]
the	↔ [D, + definite]
Ø	↔ [D]

An important aspect of the Late Insertion Model is that when two vocabulary items have features that are non-distinct from the same terminal node, the article with the greater number of features is inserted. In the above diagram, Ø is potentially insertable in the same contexts as *a* and *the*. However, in the relevant cases, *a* and *the* have a greater number of matching features and thus block the insertion of Ø (Lumsden, 1987, cited in Hawkins et al., 2006). Hawkins et al. (2006)

used this approach to explain how individual Japanese speakers included specificity as a key feature, which led them to make incorrect article choices in certain contexts.

Another explanatory framework that has been suggested to account for variability in the use of functional morphology is the Missing Surface Inflection Hypothesis (Prevost & White, 2000), according to which "it is at the surface morphological level that inflection is assumed to be absent, rather than at the abstract featural level" (p. 108). Under this hypothesis, variable production of L2 features is attributed to difficulties in mapping particular functional features to their surface morphological forms. The Missing Surface Inflection Hypothesis (MSIH) was proposed in opposition to the Impaired Representation Hypothesis (IRH) which argues that L2 inflection is essentially impaired, due to lack of functional categories or features (Eubank, 1993, cited in Prevost & White, 2000). The MSIH proposed that L2 learners indeed have unconscious knowledge of the functional features but have difficulties with the "overt realization of morphology" (p.104).

With the goal of finding support for the MSIH, White (2003) tested the predictions based on the MSIH by examining verbal and nominal inflection as well as relevant syntactic properties of a fossilized endstate grammar of L2 English of an adult native Turkish speaker. According to the MSIH, it was hypothesized that "if variable suppliance of inflection is found, and if presence or absence of overt morphology in the L1 plays a role, then a native speaker of Turkish should in principle show greater accuracy on inflectional morphology" (p. 133). With regard to the use of determiners, White expected less accurate performance on the determiner system since Turkish does not have determiners. The participant was a 50-year-old female at the time of initial data collection. Four interviews were conducted with the participant over a two-month period to gather spontaneous production data. After an 18-month period, a fifth interview was carried out

to determine whether the participant's grammar could be characterized as being in a steady state.

The results showed that the use of verbal inflections were generally accurate while there were some instances of omission of tense and agreement morphology. With regard to the use of nominal inflections and function words, the participant supplied plural morphology accurately, but determiners tended to be omitted frequently. In particular, the omission of the indefinite article was most frequent. There was a statistically significant difference between suppliance of definite versus indefinite articles. In this regard, White (2003) suggested that less consistent performance on determiners might be an L1 transfer effect.

Further, White (2003) compared these results with previous research which suggested that L2 learners of English tend to associate the definite article with [+ Specific] and the indefinite article with [- Specific], hence overusing the definite article in specific indefinite contexts (i.e., Huebner, 1979; Thomas, 1989; Young, 1996). However, the participant never used definite articles in place of indefinites or vice versa. Thus, White (2003) concluded that the Turkish participant knew that the English article system does not encode +/- specific distinction. Rather, the participant's production data showed that she was aware of +/- definite distinction. A written elicited production task as well as grammaticality judgment task, which were conducted together with oral production task, suggested that the participant was very sensitive to constraints on the use of definite and article articles, even though she frequently failed to produce them in spontaneous conversation. Observing discrepancy between the subject's knowledge and performance, White (2003) pointed out that "one must distinguish between abstract features, such as tense, agreement and definiteness, and how they happen to be realized or spelled out morphologically" (p. 139). It was also argued that the data supported the Missing Surface Infection Hypothesis which claims that L2 grammar of learners may contain abstract functional

categories and structures but such categories and structures may not be directly expressed in the surface structure. Therefore, White (2003) suggested that the missing inflection is "in some sense a surface phenomenon" (p. 139), rather than impairment to interlanguage grammar. However, White's findings can be interpreted differently as a result of L1 transfer effect, which would suggest that learners do not have the functional morphology of articles in English. This concern needs to be further clarified.

Trenkic (2007) contended with the MSIH and the Prosodic Transfer Hypothesis (PTH). The PTH suggests that a transfer of L1 phonological representations might interfere with the production of L2 morphology (Goad and White, 2004). Based on the production data of advanced Turkish learners of English, Goad and White (2004) argued that some functional morphemes such as articles in L2 may be deleted in production if certain prosodic structures required for presenting such morphemes are not available. In order to substantiate that both the MSIH and the PTH are not appropriate to account for the tendency of article omissions, Trenkic (2007) examined article production by Serbian learners of English and argued that the tendency of these learners to omit articles more frequently in adjectivally pre-modified (Art + Adj + N) contexts than in non-modified contexts could not be adequately accounted for by these hypotheses. Instead, Trenkic claimed that the error patterns produced by Serbian learners of English were related to the learners' misanalysis of the English article system as well as cognitive processing limitations.

Trenkic's (2007) study included 48 adolescent and adult Serbian learners of English to determine whether higher omission rates in Art + Adj + N contexts extend to written production. Serbian is also a language that does not have an article system. The participants carried out two tasks:

(a) a map task: a communicative task in which two participants exchange information which is only partially shared, and

(b) a written translation task: participants were given a written text in Serbian to translate into English.

The results obtained showed an asymmetry in article production in adjectivally modified and non-modified contexts in both oral production and written production, thereby confirming that the Prosodic Transfer Hypothesis cannot account for this tendency. Trenkic (2007) held that L2 learners whose L1 does not grammaticalize definiteness may analyze English articles as nominal modifiers and treat them in production accordingly, which is dubbed the syntactic misanalysis account. in other words, "their article production is lexically based (i.e., articles are treated as lexical forms) and pragmatically motivated (i.e., motivated by the perceived need to express the meaning they encode for the learner)" (p. 315). Also, "given the production of misanalyzed functional elements critically depends on the stage of production that is open to general cognition (conceptualization), it will be constrained by the available attentional resources" (p. 321). Therefore, learners may omit articles more frequently in the adjectivally modified contexts due to cognitive burdens of expressing two pre-nominal modifiers.

In a more recent study, Trenkic (2008) tested the syntactic misanalysis account with Mandarin learners of English by examining the substitution errors (i.e., *the* is supplied in indefinite and *a* in definite contexts). Trenkic (2007) assumed "a grammatical deficit, suggesting that L2ers from article-less L1 backgrounds persist in misanalyzing English articles as adjectives" (p. 2). Therefore, their article choices are governed by lexical and pragmatic considerations, and substitution errors may reflect the meanings that learners attribute to English articles. According to Trenkic (2008), with the influence of the power of the labels "definite" and "indefinite" articles, L2 learners may attach the following lexical meanings to English articles

59

misanalyzed as adjectives:

> *the*, adj, definite (that can be identified)

> *a*, adj, indefinite (that cannot be identified) (p. 15).

Lexical or referential meanings attributed to articles are based on objective identifiability, rather than discourse identifiability, the procedural meanings ARTICLES proper denote. If the speaker is familiar with identifying attributes of a referent, he may incorrectly use the definite article in indefinite article contexts where the hearer may not receive sufficient contextual information to identify the referent.

Trenkic's (2007, 2008) proposal that L2 learners of English incorrectly analyze articles as adjectives is insightful and supported by his data. However, in light of interlanguage development, such misanalysis needs to be accepted and understood as part of language acquisition. To call this type of misanalysis a deficit seems inaccurate because second language learners are known to develop new hypotheses about target language features, test them in communicative situations and revise them constantly in their language learning process.

Yoon (1993) examined the perception of noun countability by native speakers of English and Japanese speakers of English in order to investigate a possible link between noun countability and article acquisition. Traditionally grammar textbooks treat the indefinite article and the definite article under a single category of "article". However, Yoon argued that the issue of quantity and countability of nouns is only relevant to the use of the indefinite article. The definite article can be used for both count and non-count nouns. Furthermore, Japanese does not have a similar article system as well as grammatical features that encode noun number and countability. In this context, the researcher tested a hypothesis that intuitive judgment on noun countability by Japanese speakers of English would affect the use of indefinite vs. zero articles in

particular contexts.

For the study, 27 Japanese graduate students studying in the U.S. and 31 native English speakers who were in the ESL teacher training program participated in the test. The first part of the test was designed to evoke participants' intuitions regarding countability of 87 nouns without context, which were selected from two essays written by native English speakers. The second part of the test was a cloze procedure on the original essays to ask the participants to supply indefinite, definite and zero articles.

As for the first part of the test, the study revealed that not all speakers of the same language perceive noun countability in the same way. For instance, only 20% of the nouns were perceived as countable by all of the English native speakers. Also, more nouns were perceived to be countable than uncountable. Countability perception of certain nouns was reversed, where some non-countable nouns were considered countable by the Japanese participants and vice versa. This finding is indicative of the fact that certain nouns may be countable in some languages and non-countable in others. The analysis of the second part of the test was concerned with indefinite article suppliance in obligatory contexts (SOC). In the cloze test, there were 24 obligatory indefinite contexts. While 95% of the native English speakers supplied the indefinite article, only 61% of the Japanese participants used the indefinite article appropriately. With respect to the nouns that the Japanese speakers identified as countable, the study reported that the Japanese speakers did not have any problem of inserting the indefinite article in the cloze procedure. However, when the participants identified countable nouns inaccurately as uncountable, they tended to commit article errors. In fact, 89% of the Japanese participants chose the zero article for nouns that they judged as uncountable, regardless of the given contexts. Yoon (1993) thus concluded that further investigations of noun countability perception bear theoretical

and pedagogical importance.

Yoon's study made a significant contribution by investigating the topic of countability perception, which was not considered as a variable in the previous research. Yoon's work showed that perceived countability may influence the choice of articles by L2 learners whose L1 encodes noun numbers in a different way than English. However, it is necessary to question the validity of the underlying assumption that English nouns can be divided into the fixed classes of countable and uncountable nouns. In fact, the majority of uncountable nouns can also be used for countable reference. For instance, *water* can be used with zero article to refer to substance as in *Three quarters of human body are made up of water*. Or one can use a singular indefinite article to refer to a bottle of water as in *I bought a water*. This word can also be used in a plural form as in *US regulators vote to ban commercial fishing in Arctic waters*. Therefore, countability as a fixed attribute of English nouns should be questioned.

In another study on L2 use of English articles by learners whose L1s do not have an article system, Ekiert (2010) argued that the difficulty that English articles pose for L1 Slavic learners is associated with conceptual restructuring from the L1 to the L2. In particular, Ekiert (2010) investigated if and how learners who do not formally mark definiteness in their L1s express definiteness and its meanings when they use articles in L2 English. A mixed-methods research design was employed, in which 63 adult ESL learners, native speakers of Slavic languages, carried out several written tasks as well as post-task verbal reports. The results from a variety of quantitative and qualitative analyses revealed that different article uses posed unequal challenge to the participants. Specifically, article use in first-mention contexts for specific referents was shown to be particularly problematic. The results indicated that the participants' article use was often conditioned by speaker-centric assumptions about the status of referents.

Ekiert (2010) suggested that "the notion of familiarity, understood as the assessment of what is *known to the hearer*, may need to be isolated as one of the potential suspects responsible for the difficulty that English articles pose for L1 Slavic learners" (p. 2). In summary, Ekiert's study reported that speakers of L1s that do not have an article system may experience difficulty in learning to use English articles because the analysis involves the restructuring of conceptual distinctions with regard to the notion of definiteness and its instantiation in the form of articles.

This section reviewed a variety of approaches that scholars have taken to provide explanations for the difficulties associated with L2 acquisition of English articles. The findings from previous research are useful in interpreting the data on implicit and explicit knowledge of articles by Korean learners of English in this current study. In the next section, we review previous research on how to teach articles to L2 learners.

2.3.3. Pedagogy of English Articles.

Following the previous section that discussed research on the use and acquisition of English articles by L2 learners of English, this section provides a review of literature on the teaching of articles. As discussed extensively in the previous sections, it has been well documented that L2 learners of English, especially when their L1s do not employ a similar grammatical system, face difficulties in acquiring proper use of articles. For this reason, several pedagogical proposals and recommendations have been made to assist ESL teachers and learners alike for successful acquisition of English articles.

In an attempt to explore whether ESL textbooks reflect actual language usage of English articles, Pica (1983) examined patterns of article use by native speakers of English in various speech events and compared them with commonly cited rules for article usage in textbooks. The

first type of speech events was requesting and giving directions. The researcher collected data on article use in this type of situation by asking people on the street for directions. The study revealed several important findings. The pattern of using *a* for introduction and *the* for second mention was not observed in the corpus. Native English speakers frequently used *the* for the first mention, and used pronouns, synonyms or proper nouns for second mention. In this regard, Pica (1983) pointed out that the use of *the* with first-mention items in relations of association, synonymy and entailment was a frequently observed pattern in the corpus but was not made explicit in the grammar texts.

The second type of speech events was ordering food at restaurants, which included a variety ranging from expensive restaurants to fast food chains. The analysis of the speech event data produced by native speakers of English uncovered that food items were introduced with *the* only in the expensive and moderate priced restaurant categories. The zero article occurred only in moderate lunch restaurants, truck stands and fast food chains. The indefinite article *a* was used in all types of restaurants. Based on the patterns of article use emerged in the corpus, Pica (1983) concluded that the linguistic information required for appropriate interpretation of articles is often discourse-related, which makes it difficult to master the article system through textbooks and classroom practices. Article use rules that appeared in instructional textbooks implied that article choice was fixed "once linguistic and situational context were identified" (p. 230). These rules did not reflect discourse variables that affected article choices by native English speaking participants.

Based on the findings described above, Pica (1983) proposed the following recommendations for teaching English articles. First, classroom practices such as ordering food can be a non-frustrating lesson for beginner learners as article choice did not seem to influence

communicative effectiveness in ordering food. Second, students should be made aware that the use of *a* and *the* for introduction and second mention of an item is not as common a pattern in actual discourse. Thus, learners should be advised to consider relevant discursive and contextual factors instead of applying textbook rules in a rigid manner. Third, it seems that mutual identifiability of a referent influences the use of *the*. Therefore, with regard to giving directions, students should be told to use *the* with a qualifying description if they are not certain of their addressees' previous experiences with the target location. Finally, the classroom cannot provide a variety of language experiences necessary for exposing the vastness and complexity of the article system. Therefore, Pica (1983) underscored the importance of encouraging students to interact with English speakers in numerous, diverse, and frequent communicative situations.

Pica's research is significant in that it showed that the textbook grammars are not effective pedagogical tools since they do not reflect actual usage of English speakers in several contexts. In particular, her data revealed that one of the most extensively taught rules for article use in English – namely *a* for first mention and *the* for second mention – is not always followed by native speakers of English but is heavily influenced by other discourse factors. This finding suggests that learners of English need to be taught to pay attention to various discourse factors to determine which articles to use in a given conservational context. Accordingly, Pica (1983) shifted away from a traditional grammar-based language instruction to a more communicative approach in which interaction in various communicative situations affords learners with opportunities to develop sensitivity to discourse factors that influence correct article choices.

Master (1990) offered a binary schema of English articles as an approach to teaching articles to learners of English. Recognizing the usefulness of providing a one-form one-function correspondence to language learners, Master proposed that the zero article is used to classify a

noun and *the* to identify it. As for the indefinite article, he stated: "as *a* is derived from the word

one and therefore only applies to singular countable nouns, it is considered a variant form of zero

rather than a separate category of articles" (p. 465). The binary schema for article choice based

on classification and identification is presented below (p. 470).

Figure 2. 3. The binary classification of English articles
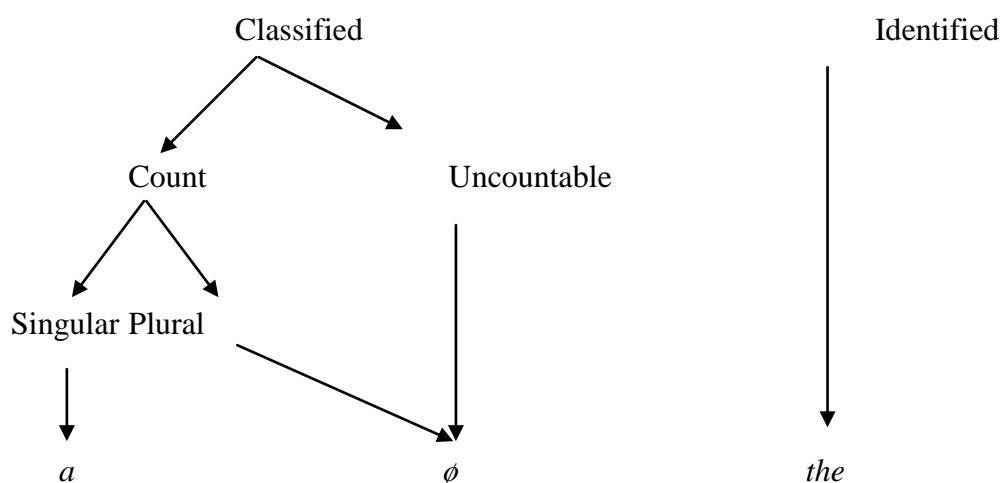
After presenting the binary system, Master provided various aspects of classification and

identification, which reflect instructional sequence as well (p. 471).

Classification (a, ø)	Identification (the)
Count/uncountable	
First mention	Subsequent mention
	Ranking adjectives
	Shared knowledge
Defining post-modification	Limiting post-modification
Partitive *of*-phrase	Descriptive *of*-phrase
Intentional vagueness	
General characteristics	
Existential *there* and *it*	
	Generic *the*
Classified proper nouns	Proper nouns (ø and *the*)
Idiomatic phrases	Idiomatic phrases

According to Master (1990), the instruction of articles begins with count vs. uncountable

distinction. When learners are able to choose between *zero* and *a* based on countability, the next

distinction in the above chart is introduced, which is a contrast between first mention and subsequent mention. The instructional sequence ends with the introduction of idiomatic phrases.

There are limitations to the approach Master (1990) proposed. First, the notions of classification and identification are not clearly defined. The model provides only descriptive aspects of the zero article and the definite article. Without clear definitions and support for the contrast between classification and identification, learners of English will be confused. Secondly, his proposal of treating *a* as a variant form of zero for singular countable nouns is not adequately supported. Other researchers (Christophersen, 1939; Hawkins, 1978; Reid, 1991) have recognized that contrast exists between any article and no article. For instance, Christophersen argued that any article actualizes an abstract concept into an actual entity with precise limits. Teaching *a* as a singular form of zero runs the risk of ignoring important differences between *a* and zero and misleading learners to believe that the choice between *a* and *zero* is simply based on singular and plural nouns. Finally, Master's suggestion to teach the count vs. uncountable distinction to beginning learners is problematic as nouns in English are not inherently distinguished for countability (Reid, 1991). For example, even a so-called 'mass' noun can be used as a singular noun with the indefinite article *a*, as in *I ordered a coffee and a bagel*. Learners of English need to be aware that many of traditionally termed uncountable nouns are in fact used in singular and plural forms in various discourse contexts.

Shifting away from his original binary system where the zero article was made contrast with the definite article, Master (1997) adopted Chesterman's (1991) analysis that there are two types of zero: the *zero* article and the *null* article. According to this analysis, the zero article is the most indefinite and its function is to "remove the boundaries that make nouns discrete" (p. 222). For instance, the zero article contrasts with the indefinite article with regard to the

distinction between count and mass nouns (*The boys ate chicken.* vs. *The boys ate a chicken.*). In other words, the indefinite article defines a boundary that makes a formless entity discrete and countable. As for use, the zero article is employed in such contexts as first mention, general characteristics, existential *there* and defining post-modification. On the other hand, the null article is the most definite and its function is to name a one-member set. The contexts for the use of the null article include proper names and familiar reference (e.g. *I went to bed late.*). The opposition of null is the definite article, which identifies a referent or indicates that "the speaker either presumes a noun to be singled out and identified for the hearer or instructs the hearer to do so" (p. 225). With regard to the frequency of English articles, Master examined five written genres, which included research journal, science magazine, news magazine, novel and play, and found that zero, which combines the zero and the null articles, is the most frequently occurring article, followed by *the* and *a*.

Following these discussions, Master (1997) recommended teaching approaches based on language proficiency. For learners with beginner proficiency, Master suggested the use of photographs or actual items to teach countability of nouns and the introduction of mass nouns at a later stage. Master recommended more cognitive methods for intermediate proficiency learners, such as Master's (1990) binary system. He held that "successful learning is more likely to occur if sufficient time is spent on practicing a single distinction at a time until students feel relatively comfortable with it" (p. 226). For advanced learners, Master (1990) argued that rules may not be useful and suggested that a lexical rather than a syntactic approach to article pedagogy is appropriate. Also, he recommended that "students should be encouraged to keep records of their errors so that they become in essence researchers on their own linguistic behavior" (p. 227).

Even though Master (1997) expanded his previous framework (Master, 1990) to include

the distinction between zero and null articles and provided pedagogical suggestions based on the frequency of articles in written language, there are still questions about the validity of his approach. Master (1997) posited nearly opposite meanings for the zero and the null articles when, in fact, they are manifested in the same way on the surface as the absence of overt article. As several researchers (Christophersen, 1939; Hawkins, 1978; Reid, 1991) point out, they may not be separate types of zero article but need to be conceptualized as a single form. Also, his survey of frequency of the English articles was limited to written genres. However, patterns of article use may differ according to speech genres. Thus, it is important to conduct a thorough survey of the use of the English articles across various written and spoken genres of language in order to help English language learners understand the article system in a comprehensive manner.

More recently, Master (2002) applied the concept of information structure to the pedagogy of English articles. In the canonical information structure, given information occurs in the subject position and new information in the predicate. With regard to the article system, given information is marked with *the* and new information is marked with *a* or *zero*. Master examined a sample text of two passages in order to examine to what extent the canonical information structure is obeyed. Given information was found to follow the canonical structure most of the time while violations were found in more than a half of the instances for marking new information. Master (2002) held that "since the canonical structure is obeyed in roughly two thirds of the total noun phrases in the two passages, it would be a potentially useful pedagogical tactic to encourage NNS students to use canonical information structure as a preliminary framework in choosing the correct article" (p. 340).

Master (2002) tested the effectiveness of the use of information structure in teaching the article system by conducting a study in which 48 adult ESL learners at an intensive ESL institute

were divided into three groups: One group of English language learners was taught the article system using information structure, another was taught using a traditional explanation, and a third did not receive any instruction. The results of the study showed that the group taught with the information structure framework made "small but noticeable improvement" (p. 331) compared to the other groups. Based on this finding, Master (2002) suggested that teachers explain that noun phrases in the subject position are marked with the definite article whereas those occurring to the right of the verb are marked with the indefinite article *a* or the *zero* article. Master also argued that this approach would help learners to correctly choose articles in the beginning stages of English language development.

In another study that focused on instruction of English articles, Akakura (2009) applied a cognitive view of language acquisition and studied the role of explicit language knowledge in L2 acquisition on implicit knowledge. Specifically, the scholar examined the effect of explicit instruction on implicit knowledge of English articles. To this end, a Web-based computer program was designed to provide lessons on generic and non-generic articles to 49 adult L2 users with some prior knowledge. The participants took the pre-tests before the instructions and the post-tests after the instructions. In order to verify the effect of explicit instructions on articles, a control group of 45 adult L2 learners also participated in the pre- and post-tests but they did not receive explicit lessons. The pre- and post-tests consisted of two oral production tasks designed to measure implicit knowledge, which included elicited production and spontaneous production tests, and two written tasks designed to measure explicit knowledge, which were grammaticality judgment and metalinguistic knowledge tests.

Results obtained from the study revealed that the participants achieved gains on the post-tests and these gains were sustained six weeks after instruction. Moreover, gains on implicit

knowledge continued to increase in the delayed post-test. Pedagogically, these results suggested that explicit instruction on English articles based on a cognitive grammar approach may be beneficial to English L2 users. In particular, the study revealed that the explicit instruction may promote the implicit knowledge of article use.

As discussed above, difficulties in teaching articles to L2 learners of English stem from a variety of sources. There has not been an effective pedagogical model for teachers to use in their classrooms. Also, as Pica (1983) maintains, there seems to be discrepancy between textbook rules on article use and the actual article use by native speakers of English. As a result, learners may receive mixed input with regard to article use. In other words, textbooks may suggest certain article usage rules, while learners may find that the ways native speakers use articles do not always follow those textbook rules. This discrepancy may cause confusion and frustration on the past of L2 learners of English when they try to find patterns of article use and develop learner hypotheses about how articles are employed in English.

To account for these pedagogical challenges, White (2010) conducted a research project with three main objectives in mind: (1) to identify how articles are currently explained by ESL textbooks and teachers, (2) to propose a systematic perspective through which to interpret article meaning, and (3) to examine how exposure to this new perspective influences the ways international MA TESOL students explain articles when their L1s do not mark definiteness with articles.

To achieve the first objective, White reviewed two leading ESL grammar books and one article workbook to investigate how they present articles to readers. The author found that the descriptions found in textbooks do not seem to cover the entire range of article uses. Rules often seem to be based on collocations rather than conceptual meanings that the speaker intends to

convey. Following a review of textbooks, an experiment was carried out in which twelve ESL teachers were asked to write explanations for twenty examples of article use found in authentic texts. The survey found that teachers employed similar terminology found in textbooks, and since textbooks fail to provide a coherent and unified account of English articles, learners may receive different explanations for the same article use.

Secondly, White proposed a conceptual framework that offered a unified account of the articles in English. Following Langacker (2008b), the author argued that articles are symbolic tools and presents a schematic imagery "as a means toward interpreting the influence articles have on meaning construction" (p. 54). In this framework, noun phrases headed by *the* prompt conceptualizations that are situated within the *discourse frame*, which include situation frame, text frame, and concept frame. The discourse frame is an abstract conceptual space to which the speaker and the hearer have access and evoked through the use of the definite article. Conceptualizations of *a* and unstressed *some* are explained in terms of individuation. In turn, all uses of the zero article prompt a construal of unboundedness, which is equivalent to having no limits within the conceptual schema. When no article appears before an NP, the referent is conceptualized as unbounded, or without limits.

Finally, White presented this conceptual framework of English articles to five MA TESOL students with L1s of Korean, Thai, and Chinese through a series of training sessions for the purpose of determining the effect of the framework on article instruction. The participants' explanations of examples of article use in authentic texts before and after the exposure to the framework were analyzed for any changes. White (2010) found that explanations after exposure to the framework were more unified across individual article uses. Thus, the author concluded that the conceptual framework of the articles may help teachers to structure their article

instruction in a more systematic and concrete manner.

2.3.4. Summary.

In this section, a review of the literature on the acquisition and pedagogy of English articles was provided. Research on L2 acquisition of articles has been extensive and undertaken from multiple perspectives. The current study has been informed by relevant research findings with regard to data analysis and the interpretation of the results.

Several descriptive studies on L2 acquisition of articles documented certain systematicity in the variable production of articles by L2 learners of English. The first finding that informed the current study is the observed directionality in article acquisition among learners of English whose L1s do not have an article system. As indicated in a longitudinal study conducted by Parrish (1987) with a Japanese learner of English, the definite article seems to be acquired before the indefinite article. Thomas (1989) also found that the correct use of the indefinite article was delayed. As English language learners start to use *the* more frequently and more accurately, they have been found to go through a period of overgeneralization of the definite article (e.g. Hawkins et al., 2006; Ionin et al., 2004; Liu & Gleason, 2002; Parrish, 1987; Thomas, 1989; Zegarac, 2004). The current study investigated whether the Korean participants exhibited similar patterns of article use.

Secondly, several scholars provided descriptive accounts of the learners' tendency to omit articles in obligatory contexts in both speaking and writing with varying frequencies (Butler, 2002; Ekiert, 2010; Mizuno, 1999; Park, 2006; Parrish, 1987; Robertson, 2000; Yoon, 1993; Young, 1996, among others). In particular, Park (2006) argued in his study that the rates of article omission did not correlate with language proficiency levels in the written production of articles

by Korean learners of English. In the current study, omission of the two articles was examined with regard to production discourse (i.e., spoken and written production) and proficiency levels.

Finally, the current study examined the influence of the presence of adjectives in NPs on article accuracy reported by Trenkic (2007, 2008) As reviewed in Section 2.3.2, Trenkic found higher rates of article omission when nouns were modified by adjectives. Based on these observations, Trenkic (2007) proposed a syntactic misanalysis account that claims that L2 learners of English from L1s without a similar article system tend to mis-analyze English articles as nominal modifiers and treat them as lexical items. Therefore, it was argued that the accuracy rate of article usage decreased with the presence of adjectival modifications due to the cognitive burden of expressing two pre-noun modifiers. In the current study, the variable of the presence of adjectives in nominal phrases was examined to determine if the syntactic misanalysis hypothesis was applicable to a group of Korean participants in speaking and writing.

In addition to documenting prevalent tendencies in article use by L2 learners of English from Korea informed by a body of literature, the study aimed to understand reasons behind the observed variability in article use. Previously, several explanatory hypotheses have been proposed to account for the apparent challenges associated with L2 acquisition of articles, which were reviewed in detail in Section 2.3.3. To gain insight from the learners to understand the sources of difficulties, Butler (2002) attempted to document learner hypotheses on English articles reported by the Japanese learners of English. The current study utilized a concurrent think-aloud protocol to understand the metalinguistic thinking processes behind article choices.

As discussed in this section, considerable progress has been made in the field of L2 article research, but the article system is still considered one of the most difficult grammatical features to acquire, especially to learners whose L1s do not have a similar article system. While

the current study has been informed by previous research, it also intends to address several gaps

identified in the literature review. The purpose of the study is to contribute to the field by

expanding our understanding of how learners use and acquire English articles. To respond to this

objective, several research questions were established to guide the inquiries made in the current

study.

2.4. Research Questions

The present study seeks to explore implicit and explicit understanding of English articles

by Korean L2 speakers of English. In light of the previous research discussed above, the present

study intended to contribute to the existing research literature on L2 acquisition of articles by

providing a comprehensive account of how learners of English understand and use articles in

various contexts when their L1s do not make use of articles. As introduced in Chapter I, the

specific research questions motivating the present study are the following:

1. What patterns in the use of articles are evidenced in Korean L2 speakers of English at

   distinct levels of English language proficiency?

2. How are these patterns distinct in speaking and writing?

3. What kinds of metalinguistic awareness do Korean L2 speakers of English have at

   different English language proficiency levels?

4. How are hypotheses about the use of English articles of Korean L2 speakers of English

   different from those of native English speakers?

5. How are Korean L2 English speakers' hypotheses about English articles operationalized

   in their speech?

The first two research questions are addressed in Chapter IV as they pertain to article use in

spontaneous production. The last three questions are discussed in Chapter V with the focus on explicit and metalinguistic understanding of English articles. Before moving to the results chapters, the next chapter describes the methodological aspects of the current study to illustrate how the study was designed to effectively respond to the research questions.

Chapter III. Methodology

In this chapter, the methodological procedures of the study are described. First, the design of the study is discussed. This is followed by an introduction of the study's participants and the data collection method. The chapter ends with an explanation of how the data gathered from various data collection instruments were analyzed quantitatively and qualitatively.

3.1. Design

The present study is a cross-sectional study that employed a mixed-methods design. It is cross-sectional since the study participants consisted of L2 learners at different levels of proficiency. A typical longitudinal study in the field of SLA investigates interlanguage development of a particular L2 feature of a small number of participants for an extended period of time. Compared to a longitudinal study, a cross-sectional design is advantageous in that it can study a larger number of participants efficiently in a short period of time. Since the goal of this study is to provide a comprehensive picture of how Korean learners of English use and understand articles, a cross-sectional study was deemed appropriate and necessary.

The present study employed mixed-methods to analyze data. That is, it used quantitative and qualitative data analyses. A mixed-methods study is usually defined as a research design in which the researcher combines quantitative and qualitative research techniques, methods, and

approaches to data collection and analysis into a single study and it provides more comprehensive evidence for the research problem under investigation. This type of research design offsets weaknesses of a study that is either exclusively quantitative or purely qualitative.

As this study aimed to investigate implicit and explicit knowledge of English articles by Korean learners of English, careful consideration was made in designing the research methodology. As Akakura (2009) points out, there are currently no pure measures of either implicit or explicit knowledge. However, recent empirical studies in assessing knowledge of L2 helped advanced our understanding on various test measures that can fairly accurately discriminate between implicit and explicit knowledge. For instance, Han and Ellis (1998) found that it is possible to measure implicit and explicit L2 grammatical knowledge separately. In their study that investigated four types of tasks (i.e., an oral production, timed grammaticality judgment test, un-timed grammaticality judgment test, and an interview eliciting verbalized metalinguistic knowledge), the scholars concluded that the amount of task processing time was a discriminating factor between implicit and explicit knowledge.

Among various tasks that measure implicit and explicit knowledge of a second language, an oral narrative task, a written narrative task, a grammaticality judgment task, and a think-aloud task were selected for this study to enable triangulation. The first two tasks are considered measures of implicit knowledge of English articles and the last two tasks, measures of explicit knowledge. In the following, the characteristics of each task are discussed.

Free constructed responses such as oral narratives require the production of language with a communicative purpose. Thus, free production tasks require a language user to focus on meaning rather than form. As implicit knowledge of language is manifested in spontaneous language use, free constructed responses are most appropriate in evaluating implicit knowledge.

Among various types of spontaneous language use, a story can provide a realistic context for grammatical forms while engaging in meaning.

This study employed an oral narrative task and a written narrative task to measure implicit knowledge of English articles by Korean learners of English at distinct proficiency levels. This design is based on the recognition of the benefits of a narrative task in gathering L2 language production data. A narrative task is a commonly used spontaneous language production task. According to a general definition, a narrative is a text type that recounts and organizes actions into a unified whole that has meaning. Narratives are the central means by which people make sense of their experiences. Also, the narrative discourse is believed to be relatively familiar to adult language users. Therefore, Ellis (2003) assumes that a lowered cognitive load is involved in a narrative task completion. For these reasons, an oral and written narrative task was chosen as an instrument to obtain spontaneous production data from the participants.

In obtaining narrative data from the participants, two kinds of discourses were distinguished: spoken and written discourses. The spoken discourse was elicited in this current study because it is considered a more accurate measure of implicit knowledge of language. For instance, Ellis (2005) pointed out that spoken production is better suited for the purposes of assessing implicit knowledge because speech is produced spontaneously while writing allows more time to possibly monitor production. Meanwhile, the written discourse was obtained to investigate the effect of monitoring of language production on article use. It is generally expected that grammatical accuracy is higher in writing than in speaking, but this hypothesis has not been empirically tested with regard to English articles. Since the participants were not informed that articles were a target linguistic feature of the study, the amount of attention they directed to accurate production of articles would be dependent on the level of sensitivity and awareness they

have about the article system.

With regard to the measures of explicit knowledge, two tasks were employed in this study: a think-aloud task and a grammaticality judgment task. The reason for the use of two metalinguistic exercises is related to the notion of explicit knowledge (Ellis, 2004). After reviewing several studies that explored conscious knowledge of particular L2 language features, Ellis (2004) proposed "L2 explicit knowledge" (p. 228) as a theoretical construct and defined it as "the declarative and often anomalous knowledge of the phonological, lexical, grammatical, pragmatic, and socio-critical features of an L2 together with the metalanguage for labeling this knowledge" (p. 244-245). In other words, L2 explicit knowledge can be expressed in terms of the conscious knowledge and the metalanguage. As summarized by Ellis (2004), conscious knowledge of L2 language features has been traditionally measured by various types of grammaticality judgment tasks, while the L2 metalanguage of learners has been measured through verbal reports and interviews. Therefore, this study employed a think-aloud task and a grammaticality judgment task. The characteristics of each task are discussed in what follows.

A think-aloud task is understood as a type of verbal reports in SLA. In general, verbal protocols are used to gain insight into explicit knowledge either introspectively during a task or retrospectively. For instance, Butler (2002) adopted a retrospective approach and interviewed individual participants to verbally explain their article choices after completing a fill-in-the-blank task. Similarly, Ekiert (2010) employed a retrospective recall task, in which participants were asked to provide reasons for article choices after they completed the missing word task. As Akakura (2009) summarizes, retrospective reports are advantageous because the actual task is not compromised and there is less intimidation to perform. However, limitations of retrospective recalls include concerns that self-reports may have problems with reliability and participants may

have forgotten actual reasons for their decisions. They may even attempt to provide post-hoc rationalization of their actions, which may be different from the real reasons for their actions. In order to avoid these concerns, an introspective verbal protocol was chosen for this study, which is referred to as the think-aloud task in this study.

Concurrent reports, such as think-alouds, generally involve verbally explaining the reasons for choosing one particular grammatical structure over another as one works on a task in a monologue. In this regard, Roehr (2006) stated that a think-aloud has advantages of potentially allowing direct access to thought processes during a task with "minimum of forgetting and post-hoc rationalizations of actions or decisions on the part of the participant" (p. 181). It is worth mentioning that one of the disadvantages of concurrent verbal reports is that L2 users with lower English language proficiency may experience difficulties in articulating reasons for their grammatical choices in L2. For this reason and to avoid a gap that could exist with lower proficiency participants, the participants in this study were allowed to verbalize in their native language.

The second measure of explicit knowledge of English article employed in this study is a grammaticality judgment test. The implementation of the grammaticality judgment exercise represented another way to measure conscious L2 knowledge of English articles and served to strengthen the validity of the findings from the think-aloud protocol. Han and Ellis (1998) explained that grammaticality judgment tests in SLA have been frequently administered as a receptive measure of language knowledge. Learners in these tests are asked to judge the grammaticality of decontextualized sentences. In this present study, sentences containing the article forms under investigation were provided within conversational contexts since grammaticality judgments of English article use are critically dependent on discourse contexts.

Another characteristic of the grammaticality judgment task used in this present study is that no time limit was given. Ellis (2005) pointed out that this type of task forces learners to access their explicit knowledge if there is no time pressure to complete the task. In order to ensure that participants could utilize their knowledge of articles as much as possible, they were allowed to complete the task at their own pace.

The four tasks used in the current study as data collection instruments and the measures of language proficiency (i.e., a sociolinguistic interview, a Cloze test, and a writing exercise) were tested with three learners of English from Korea – one beginner, one intermediate, and one advanced speaker of English. The three learners were not included in the actual study. The testing of the instruments was conducted to identify areas for improvement, understand the actual logistics of the implementation, and ensure the validity and reliability of the measures.

In summary, this current study was designed as a mixed-methods study that employed both qualitative and quantitative data collection instruments and allowed for triangulation of data. The four tasks used in this study are briefly illustrated here:

Figure 3. 1. Methodological design of the study

Measures of implicit knowledge   Measures of explicit knowledge
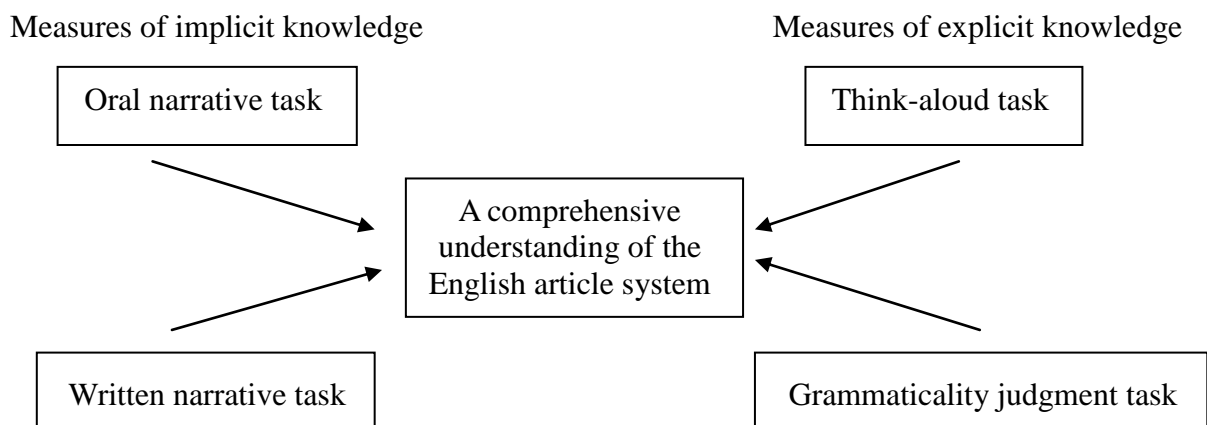
3.2. Participants

Since the purpose of this study was to investigate the knowledge of the English article system that Korean learners of English have, random sampling was not preferred. Instead, this study relied on purposeful sampling based on individual characteristics of the participants relevant to the research problem.

Participants were mainly recruited through a snowball sampling method with the use of referrals. As a doctoral student from Korea, I had access to various Korean communities in the state of New Jersey. I utilized my social networks to search for potential participants and seek their referrals to find more. Potential participants referred through the snowball method were initially screened based on several criteria. First of all, participants needed to speak Korean as their first language and have learned English as a second or foreign language. Secondly, to control a social variable of the sample, the age range of the participants was limited to their 20s. This age group had the largest pool of L2 learners of English from Korea. However, gender was not considered a variable in this study as gender has not been identified in the literature as a major social variable that influences the use and acquisition of L2 grammar features (e.g., Lee, 1997).

As a result of the snowball sampling method, 38 candidates for the study were initially recruited. Then, proficiency profiles were used to finally select ten beginners, ten intermediates, and ten advanced speakers of English. Since proficiency was a key external variable in this cross-sectional study to investigate article use and understanding, it was essential to have an accurate proficiency assessment. As such, a detailed proficiency evaluation process was employed to select a final group of 30 participants. The evaluation instruments used to screen participants for proficiency covered all four areas of language – speaking, listening, reading, and

writing – as they included three instruments: a Cloze test, a sociolinguistic interview, and a writing exercise. The following provides a discussion of each of these proficiency assessment exercises.

A key assessment tool used in this study was a Cloze test. Cloze tests have been used in several research studies on L2 article acquisition for the purpose of measuring English language proficiency (e.g., Ionin, Zubizarreta & Maldonado, 2008; Liu & Gleason, 2002). The Cloze test used in this study was designed in a similar way that prototypical Cloze tests had been administered. In other words, the first paragraph was left intact to contextualize the entire text. In the ensuing paragraphs, every tenth word was replaced with a blank. There were a total of 63 blanks to fill. The Cloze instrument used in the study is provided in Appendix A. Participants were asked to fill in these blanks with words that they believed were the most appropriate. Each participant was given 30 minutes to complete the Cloze exercise. The Cloze exercise was graded using an acceptable-word criterion, in which acceptable words in the blanks were graded as correct even though they may not be the exactly same words in the original text (Ionin et al., 2008).

As a supplementary instrument to determine English proficiency of the participants, a sociolinguistic interview was conducted. The purpose of the interview was two-fold. First, the sociolinguistic interview was used to obtain background information about each participant, including previous English language learning experiences, length of stay in the U.S., and the amount of daily English use. This approach followed the framework of sociolinguistic interviews proposed by Labov (1984). According to Labov, one of the primary goals of the sociolinguistic interview is to "obtain the full range of demographic data necessary for the analysis of sociolinguistic pattern" (p. 32) including age, residential, school, occupation and language

history. Social, educational and language background data were used to determine proficiency of the learners. In the current study, proficiency was used as a key external variable that was examined together with linguistic variables such as accuracy of article use, presence of adjectives in NPs, and categories of NP contexts.

The second purpose of the sociolinguistic interview was to assess oral language proficiency of each participant. When the participant responded to a series of questions stated in the interview protocol, spontaneous oral language was produced and audio-recorded for oral proficiency assessment. The oral proficiency evaluation was made based on the independent speaking assessment rubric informed by the TOEFL Test. The rubric is provided in Appendix B. Also, the actual protocol for the sociolinguistic interview is provided in Appendix C.

In addition to the sociolinguistic interview and the Cloze exercise, another instrument used to assess the language proficiency of the participants was a writing exercise. The writing exercise complemented the interview and the Cloze exercise with the purpose of triangulating the writing proficiency with the oral and linguistic competence to further determine the language proficiency of the participants. In the current study, the writing exercise was modeled after the independent writing section of the TOEFL Test. In the writing exercise, three essay topics selected from the official list of TOEFL essay writing topics were provided. Each participant chose one topic and wrote a short essay about the topic within the time frame of 30 minutes, which is the amount of time allowed for the independent writing portion of the TOEFL Test.

The assessment of written essays was based on the TOEFL independent writing assessment rubric, which is provided in Appendix D. The participant was provided with a laptop computer, which disabled grammar related functions. Paper and pencils were also available in case a participant wished to write on paper. Dictionaries were not allowed. The instrument given

to each participant for the writing exercise is provided in Appendix E.

In determining the language proficiency of potential participants, Cloze test scores served as a primary indicator. In addition, speaking and writing scores were used as supplementary data for a final determination of language proficiency. For instance, B9 in the beginner group exhibited a good performance in the Cloze test but received low scores in the speaking and writing evaluations. As a result, the participant was placed in the beginner group instead of the intermediate group. This multi-dimensional approach to proficiency assessment helped improve the validity of this present study. Individual participants' scores on the Cloze test, speaking exercise and writing exercise in addition to their lengths of stay in the U.S., which was gathered to determine English language exposure, and the ages of arrival are reported in Appendix F.

Group means and standard deviations of the Cloze scores, age of arrival in the U.S. and length of stay in the U.S. are summarized in Table 3.1.

Table 3. 1. Background information according to proficiency groups

Proficiency group	N	CLOZE score (perfect = 63)	Age of arrival in the U.S.	Mean length of stay in the U.S. in Years
		Mean (SD)	*Mean (SD)*	*Mean (SD)*
Beginner	10	19.9 (8.7)	23.7 (1.76)	0.95 (0.82)
Intermediate	10	31.7 (3.12)	22.2 (2.34)	2.32 (1.64)
Advanced	10	38.0 (5.18)	15.6 (3.16)	7.35 (2.40)

The beginner group was mostly composed of recent arrivals that had come to the U.S. within a year. At the time of the study, participants in this group were attending ESL classes to improve their English skills. Most of them maintained a positive attitude toward learning English but expressed frustration at the slow progress they were making in learning English. The intermediate proficiency group consisted of L2 speakers of English from Korea, who have been in the U.S. for a period of two to five years. They were taking advanced level ESL classes,

85

attending community colleges, or working at places where English skills were not critical. Most advanced level participants have lived in the U.S. for more than five years and use English extensively either at school or at work. They also had extensive contact with native English speakers and had high level of English language skills.

In addition to the 30 Korean participants in three proficiency groups, the study also used a control group of five native English speakers (NES) of Korean decent. The NES group participated in the think-aloud task and the grammaticality judgment task. The control group members were Korean Americans in their 20s. The participants were recruited from several sources on a voluntary basis, including local Korean American churches, and Korean American college associations from a large state university in the Eastern part of the U.S. All of them were born in the U.S. and had received the majority of their education in the U.S. By recruiting young Korean Americans in their 20s as the control group, the study was able to draw reliable comparisons between Korean L2 speakers of English and Korean L1 speakers of English.

3.3. Data Collection Instruments

The rationale for the design of this present study was explained in Section 3.1. To address the research questions and examine implicit and explicit knowledge of English articles by Korean learners of English, four specific tasks were designed and implemented: an oral narrative task, a written narrative task, a grammaticality judgment task, and a think-aloud task. As mentioned in Section 3.1, for the purpose of enhancing validity and reliability of the study, these data collection instruments were tested with one beginner, one intermediate and one advanced L2 speaker of English from Korean L1 background. The three learners were not included in the actual study. Final modifications were made to the instruments based on the

feedback from the testing. In the following segments, each of the instruments is described in detail.

3.3.1. Oral narrative task.

As explained in Section 3.1, the oral narrative task was carried out to obtain spontaneous speech, which demonstrates the implicit knowledge the participants have of English articles. As oral narratives are produced in real time, participants are not afforded much time to plan, correct and rephrase their language. In this sense, the oral narrative task is appropriate for examining implicit and unconscious understanding of articles.

In order to obtain spontaneous speech, a Labovian approach to collecting narratives was employed. Labov (1997) stated that "the elicitation of narratives of personal experience proved to be the most effective" (p.395) among the partial solutions to the Observer's Paradox within the face-to-face interview. As the participants in this study speak English as their second language, it might not have been easy for them to produce a large amount of language, particularly for beginner level participants. Personal stories were most likely to engage the participants and elicit most language from them. Therefore, each participant was asked to tell a personal story that was the most important or memorable to the participant. When a participant failed to produce enough language or had difficulty in producing a complete story in English, probing questions were provided. The participant's oral narrative was digitally audio-recorded for storage and analysis.

Specifically, in order to effectively elicit spontaneous oral narratives, a four-step strategy was used. First, a non-guided general question was used to ask each participant to tell a story. When the general question failed to elicit a story, the second question was a recast of the first

question by narrowing it down to an event that left a big impact on their life. The third question that could be asked to elicit a narrative was about the participant's family or friends. The participant might have felt more comfortable to talk about other people's stories. Finally, when there was still no response, a more guided question was to be asked. Most participants were able to produce oral narratives at the first step. The detailed script used to obtain oral narratives is provided in Appendix G.

3.3.2. Written narrative task.

Immediately after producing an oral narrative, each participant engaged in a writing task. Instead of providing certain topics for writing, each participant was asked to create a written narrative of the story that was already narrated orally. The reason for this design was twofold. First, as the participant had already narrated the event, the content was still fresh in their memory. Therefore, the participant did not have to spend time thinking about the content of their writing. Secondly, the researcher could assume a certain degree of control of the content of both oral and written narratives, which was useful in making comparative analysis between the two genres. However, caution was needed in obtaining written narratives because there was a concern that participants might possibly use *the* for first-mentions of referents based on the perception that these referents had already been introduced in the oral narratives. Therefore, it was emphasized that the written stories needed to be as detailed as possible, assuming that the reader had not heard of the stories. It was essential to remind them that someone else would read the story and that the reader would not know anything about the participant. This approach helped minimize any priming effect on article use when the participant wrote the story that had just been orally narrated.

For the written narrative, each participant was provided with a laptop computer, where spell check and grammar check functions were disabled. However, paper and pencils were also available in case the participant wishes to write on paper. Dictionaries were not allowed. Participants were instructed to write stories that were narrated orally and write them as detailed as possible.

As mentioned earlier, the purpose of collecting written narratives was to compare patterns of article use between writing and speaking. It is generally assumed that learners are afforded more time and attention to linguistic features in writing because writing is a more planned activity. Therefore, time limits were not placed on this task. The script used to give instructions to each participant for the written narrative task is provided in Appendix H.

3.3.3. Grammaticality judgment task.

The grammaticality judgment task required the participant to judge whether an article in a given context was acceptable or not. The grammaticality judgment task contained 16 short dialogues. No detailed scenarios were provided for each dialogue, except for a brief setting, to ensure that participants would not be primed in their judgment. In each dialogue, the final turn-taking contained a target article in an underlined bold format, as is shown in the following example:

> *At a bookstore*
> Chris: Well, I bought everything that I wanted. Are you ready to go?
> Mike: Almost. Can you please wait a few minutes? I want to talk to **<u>an</u>** owner of
> this bookstore. She is my old friend.
> OK _____   NOT OK _____

The content of the exercise was a modification of the dialogues that Ionin, Ko and Wexler (2004) employed to elicit article choices of L2 speakers of English. Of the 16 dialogues used in the

current study, ten dialogues were borrowed from Ionin et al. (2004), five of which were instances of the use of *the* and the other five were the instances of the use of *a(an)*. In addition, six original dialogues were created to include instances in which no article was required. Of the 16 instances, eight instances were constructed as appropriate uses of articles and the other eight, inappropriate. The appropriate and inappropriate instances of various forms of English articles were randomly distributed in the grammaticality judgment task, which is provided in Appendix I.

3.3.4. Think-aloud task.

In order to collect data on the metalinguistic understanding of English articles and address the research questions, a think-aloud task was used in the current study. The design of the think-aloud task was informed by Butler (2002) in the sense that the participants were given a fill-in-the-article text with blanks in places for articles and that they were asked to verbalize reasons for their article choices.

However, a major difference between Butler (2002) and the current study was that, unlike the post-task interviews with the participants in Butler's study, the participants in this study were asked to verbalize their article choices while making article selections. As admitted by Butler (2002), the post-task interview approach had limitations with regard to how clearly participants could remember exact reasons for their article choices. For this reason, this present study employed a think-aloud method to capture the thinking process at the moments of making article selections.

Another modification in the think-aloud protocol was made with regard to the kinds of texts used for the task. In Butler (2002), several short passages from different English texts were chosen for the fill-in-the-article test and only selected articles were deleted for participants to

insert. In the current study, participants received a complete written story with blanks for articles. The rationale for the use of one story was that a coherent text was assumed to provide more contextual cues for article selections that resemble authentic article use in real language situations.

Several criteria were applied in selecting the text for the fill-in-the-article test. First of all, the text needed to be in the form of a narrative in order to maintain the same discourse of language use with the oral and written narrative tasks. Second, the chosen text needed to be easy enough for L2 English speaking participants of all proficiency levels to comprehend. The fill-in-the-article text used for the think-aloud exercise was selected from a list of written texts targeted for intermediate ESL reading practices. Even though the text was chosen from reading materials for intermediate-level learners, in the testing of the think-aloud exercise, the three learners (i.e., one beginner, one intermediate, and one advanced learner) reported no difficulty in comprehending the fill-in-the-article text. The full text is provided in Appendix J.

Each participant received the selected text with blanks for articles. Each blank needed to be inserted with one of the following choices: (1) *a (an)* (2) *the*, and (3) $\emptyset$. The last choice ($\emptyset$) was used when no article form was needed. There were a total of 72 blanks, which consist of 37 instances of *the*, 20 instances of *a (an)*, and 15 instances of no article $\emptyset$.

In making the choice for each blank, each participant was asked to think out loud and verbalize reasons for each choice. As the purpose of this exercise was not to assess the English language proficiency of the Korean participants but to access their thinking and analytical processes behind article choices, the use of Korean was allowed for all participants with exception to the native speaker control group which was asked to articulate their choices in English. The think-aloud exercise was digitally recorded for data analysis.

In the next section, details regarding how the data collection was conducted regarding all four tasks.

3.4. Data Collection Procedures

The data for this study were collected over a time span of 4 months. Each participant carried out the four data collection tasks (i.e., the oral narrative task, the written narrative task, the think-aloud task, and the grammaticality judgment task) in two separate sessions. In the first session, the oral narrative and the written narrative were collected. In the second session, the think-aloud task and the grammaticality judgment task were conducted. There was a delay between the two sessions, ranging from a week to three weeks, depending on scheduling concerns with each participant. The entire data collection procedures are summarized in the following table.

Table 3. 2. Data collection procedures

	Data collection instruments
Session 1	Oral Narrative Task
	Written Narrative Task
A time gap (one to several weeks)	
Session 2	Think-aloud Task
	Grammaticality Judgment Task

The first session consisted of the oral narrative task and the written narrative task. First, a participant was asked to tell a personal story that was the most important or memorable to the participant in order to obtain natural and spontaneous speech that demonstrates the implicit knowledge of English articles. Time limits were not placed on this task. The participant's oral narrative was digitally recorded and lengths of oral narratives ranged between 10 to 20 minutes.

Immediately after producing an oral narrative, the participants engaged in the written

narrative task. Each participant was asked to produce a written narrative of the story that was already narrated orally. The participant was given a laptop to write the narrative. The laptop was provided to make it more convenient for each participant to reflect on and edit their writing. In order to ensure integrity of writing samples of the participants, the Microsoft Word features with respect to spell checks, and underlying and advising of errors were disabled. Several participants, especially beginner participants, who wished to write on paper, were provided with a pencil, an eraser, and sufficient paper. Dictionaries were not allowed. Limits were not placed on the amount of time or the length of writing. Participants were encouraged to write as much as possible and were asked to keep external readers in mind when they wrote. Each writing sample produced on a computer was saved on the hard drive of the laptop for analysis. The hand-written narratives were later transcribed by the researcher on the computer for later analysis.

In the second session, the think-aloud task and the grammaticality judgment task were carried out. Each participant first received an overview of how the think-aloud task would be conducted. The think-aloud method may have been unfamiliar to participants. Therefore, in order to familiarize them with this instrument, a brief demonstration was provided which used verb tenses as an example. In the sample demonstration, selected verbs were underlined and appropriate verb tenses needed to be selected. Possible choices were the past tense, the present tense, and the future tense. The actual script for the sample demonstration is provided in Appendix K. The questions in the sample demonstration were created especially for this current study. Furthermore, written instructions were provided at the beginning of the think-aloud task to remind the participants of the procedures.

Participants were asked to think out loud and verbalize reasons for each choice as they were making article choices for the blanks in the fill-in-the-article test. As the purpose of this

task was not to assess the English language proficiency of the Korean L2 speakers of English but to access their thinking and analysis processes behind article choices, Korean was allowed for the Korean participants. The Think-aloud task was digitally audio-recorded and transcribed for analysis.

After the completion of the Think-aloud task, each participant carried out the grammaticality judgment task. When a given article was deemed acceptable, the participant put a check mark next to *OK*. If not, the participant was asked to fix the article error in the blank next to *NOT OK*. The grammaticality judgment task was not timed to ensure that participants would take enough time to gather their thoughts and access their knowledge of English articles.

The control group of five Native English speakers participated only in the second session, which included the think-aloud task and the grammaticality judgment task. The data collection procedures for the control group were the same as the learner participants. In other words, they first carried out the think-aloud task, followed by the grammaticality judgment task.

3.5. Data Analysis Procedures

The implementation of the four data gathering tasks described in Section 3.4 produced a rich body of data regarding the understanding and use of English articles. These data were analyzed quantitatively and qualitatively to capture implicit and explicit knowledge of English articles that Korean L2 learners of English exhibited at the time of data collection. Specifically, the quantitative analyses of articles that the participants produced in their oral and written narratives focused on identifying patterns of article use in the modes of speaking and writing according to proficiency levels. The qualitative and quantitative analyses of the data from the think-aloud task were carried out to examine metalinguistic knowledge and hypotheses that the

94

participants generated in making article choices. In addition, the quantitative analyses of the data obtained from the grammaticality judgment task produced indicators of the level of explicit awareness of articles. This section provides a detailed account of how these analyses were conducted to address the research questions of this study.

3.5.1. Procedures for coding and analyzing article production data.

This section discusses how the production data from the oral and written narrative tasks were analyzed. The focus of analyses was to examine and compare patterns of article use in the oral and written production according to proficiency levels with the purpose of addressing the first research question (i.e., What patterns in the use of articles are evidenced in Korean L2 speakers of English at distinct levels of English language proficiency?) and the second research question (i.e., How are patterns of article use distinct in speaking and writing?).

Before carrying out analyses of article use in the oral and written narratives, the digital audio-recordings of the oral narratives produced by the participants were transcribed by the researcher into Microsoft Word documents. Most written narratives were produced on the computer, but several of the written narratives produced with a pen and paper were transcribed by the researcher into Microsoft Word documents.

In the analyses of oral and written narratives, the first step was to identify all the noun phrases (NPs). These noun phrases were coded for the four linguistic variables. However, proper nouns (i.e., names of people and places) and pronominal phrases (e.g., *She was excited to see her*

mother.) were excluded from the analyses.[1] The variables which were examined were the

category of determiners, accuracy of article use, presence of adjectives in NPs, and the category

of NP contexts. The data entry of nominal phrases for coding was first made in Excel

spreadsheets and then transported to SPSS for statistical analyses.

The first linguistic variable was related to the categories of determiners that appeared in

NPs. Each NP was coded for a determiner that was deemed appropriate in a given NP and a type

of determiner that was actually produced. In English, a variety of determiners can occur in a pre-

nominal position, including articles. The types of determiners used in this study for coding are

summarized in the following table.

Table 3. 3. Categories of determiners

Code	Description	Example from the data
1	Definite article (*the*)	*The* museum was very big. (A7, WN[2])
2	Indefinite article (*a/an*)[3]	We ended up not finding *a* space. (A7, WN)
3	Bare NP (i.e., no article is needed)	We found this place where it sold muffins. (A7, WN)
4	Demonstratives (e.g., *this, that, those, these*)	(…) introduced *this* wonderful Japanese restaurant. (A7, WN)
5	Possessive pronouns (e.g., *his, my, our*)	This would explain how much this drawing captured *our* eyes. (A7, WN)
6	Quantifiers (e.g., *one, two, many, a few, some, no*)	We took *many* pictures by the lake. (A7, WN)
7	Others (e.g., *another*, *each*)	She moved to *another* state and I had to give her up. (A1, WN)

---

[1] Proper nouns were not included because the use of articles in proper names does not always follow conventionalized uses of articles. Also, pronominal phrases were excluded from coding because they do not take articles.

[2] When excerpts from the corpus are presented, the proficiency group (*B* for beginner group, *I* for intermediate group, and *A* for advanced group) as well as the participant number are indicated. For example, *A5* refers to participant number 5 in the advanced group. Also, *WN* refers to written narrative while *ON* refers to oral narrative.

[3] The indefinite article has two phonological realizations in English – *a* before a consonant and *an* before a vowel. In this study, *a* is used as a cover term for both phonological representations.

After the required determiners and produced determiners for all the NPs were coded, accuracy in article use for the NPs was coded. The produced determiner was coded accurate if it was identical to the required determiner. If the produced determiner was not matched with the required determiner, such an instance was coded as inaccurate. However, it is important to note that depending on the context, there may be more than one grammatically accurate determiner choice for a given pre-nominal position. As long as a form in a pre-noun position was grammatically accurate, it was coded as accurate in this study.

Table 3. 4. Coding for accuracy in the use of articles

Code	Description	Example from the data
1	Accurate = the participant makes the best choice in the pre-nominal position.	Police officer came and he asked few questions about _the_ situation (A5, WN)
2	Inaccurate = the choice in the pre-noun slot is not the best in a given context of NP	*Police officer came and he asked few questions about the situation (A5, WN)

The next linguistic variable used in the analysis was the presence and absence of adjectives in NPs. As discussed in Chapter II, Trenkic (2007) proposed the Syntactic Misanalysis Hypothesis and argued that L2 learners of English use articles more accurately when nouns are not modified by adjectives. Lee (1997) also reported that Korean learners of English deleted articles more frequently when nouns were modified by adjectives. Thus, this variable was included to determine whether the presence of adjectives in NPs had any effect on the accurate use of the two overt articles in the oral and written narratives.

Table 3. 5. Coding for the variable of the presence of adjectives in NPs

Code	Description	Example from the data
1	Presence of adjectives in NP: when a participant used an adjective to modify a noun in a given NP	It was a totally *different* world. (A6, WN)
2	Absence of adjectives in NP: when a participant did not use an adjective to modify a noun in an give NP	We tried to look for a place to park around Fort Lee (A7, WN)

The last linguistic variable used in this study for coding was types of NP contexts. The two types of NP contexts for the definite article used in this study were informed by the classification of articles proposed by Hawkins (1978). Using the notion of a shared set, Hawkins (1978) classified usage types of the definite article into four categories: (a) anaphoric use, (b) immediate situation and visible situation use, (c) larger situation use based on shared general knowledge, and (d) associates of the referent. Among the four categories of usage of the definite article, the anaphoric use is most frequent in narratives to maintain referential coherence. Also, the rule of article use for anaphoric references (i.e., the use of *the* for subsequent-mentions of previously introduced referents) is well-known to most ESL learners as this rule is included in sections on article use in ESL textbooks. For these reasons, it was assumed that the anaphoric use of the definite article would be easier for Korean learners of English compared to the other usage. As such, NP contexts for the definite article were divided into two categories in this current study: *anaphoric definite* contexts in reference to anaphoric use by Hawkins (1978) and *associative definite* contexts that includes the other three uses of *the*. In short, the anaphoric definite contexts refer to second and subsequent mentions of a referent that has been previously introduced in the discourse, as in *I bought a book yesterday. The book was on sale.* The associative definite contexts refer to instances in which the definiteness of a noun referent is

determined by contextual factors such as shared knowledge and post-adjectival modifications, as in *I bought a book yesterday. The cover was red*.

With regard to the indefinite article, NP contexts in which the indefinite article *a* is used can be divided into two: *referential indefinite* contexts and *non-referential indefinite* contexts. Following Ekiert (2010), referential indefinite contexts are associated with the first mention of a referent that is specific yet unknown to the hearer as in *I bought a gift for you*. Non-referential indefinite contexts refer to instances in which a noun referent does not refer to any specific entity as in *I feel lonely. I need a friend*. The NP contexts for the two articles are as follows.

Table 3. 6. Categories of NP contexts in which the two articles are used

Article	Description	Example from the data
the	Anaphoric definite contexts	We went to genocide museum (..) and *the museum* well portrayed those times such as their ways of torture.. (A4, WN)
	Associative definite contexts	I and one of my friends went to the swimming pool (…) we checked in at *the front desk*.. (A5, WN)
a	Referential indefinite contexts	I wrote *a letter* to her and decided to give her later. (A1, WN)
	Non-referential indefinite contexts	At first, I thought she was *a college student*.. (A1, WN)

After the above described variables were coded by the researcher, an inter-rater reliability check was carried out to ensure that the coding performed by the researcher was reliable. The outside rater was a Ph.D. student in the language education program who is a native speaker of English. The rater's role was to code the data produced by one beginner participant, one intermediate participant and one advanced participant. The researcher gave the rater a detailed coding guideline as well as a blank coding sheet. The rater was also provided with the full written narratives as well as the transcripts of the oral narratives. The coding sheet was

already filled with noun phrases and determiners produced by each participant. The rater was asked to code each noun phrase for the variables of 'required determiner' and 'accuracy'.

After receiving the results from the rater, the rater's codings were compared with the researcher's coding. With regard to the variable "accuracy", which is the most important variable as it represents whether the article choices by the learners are matched with the target articles, there was a 94.8% agreement between the researcher and the rater ($N = 417$ agreed out of 440 NPs). To ensure a statistically reliable degree of agreement, Cronbach's α was calculated, which was .919.

After all the coding was complete, the statistical program of SPSS was used to carry out statistical analyses. For the production data, descriptive statistics as well as cross-tabulations were conducted to identify patterns of article use according to proficiency levels. A cross-tabulation yields a picture of how two variables inter-relate. This approach is useful to determine what effect an independent variable has on a dependent one, which, in this study, is the accurate use of articles. For example, the dependent variable of 'accuracy' was cross-tabulated with the independent variable of 'presence of adjectives in NPs' to examine whether the presence of adjectives in NP had effect on the accuracy with which articles were employed. Also, the dependent variable of 'accuracy' was cross-tabulated with the independent variable of 'NP contexts' to explore which NP contexts posed the greatest challenges to the participants at different proficiency levels with regard to article use. Patterns of article use according to these linguistic variables were then compared across proficiency levels, which were taken as an external variable on article use. They were also compared between the genres of oral narratives and written narratives.

3.5.2. Procedures for analyzing think-aloud data and grammaticality judgment data.

While the oral and written narratives were elicited to capture spontaneous use of articles by the participants, the think-aloud task and the grammaticality judgment task were conducted to direct the participants' attention to the grammar of English articles and evoke their explicit understanding of the article system. This section discusses how the data obtained from these two tasks were analyzed.

First, the data from the grammaticality judgment task were analyzed quantitatively using SPSS. Initially, group means and standard deviations of the task were obtained according to four proficiency groups. Then, independent samples *t*-tests were conducted to examine if there were significant differences in the mean scores of the proficiency groups. Finally, rates of accuracy for each article form as well as each NP context were examined to explore the developmental levels of the participants' explicit knowledge of articles.

The data obtained from the think-aloud task were two types: (1) the actual article choices made by the participants in the fill-in-the-article test, and (2) the reasons for such article choices verbalized by the participants. With regard to the first type of data, accuracy of article choices was coded in order to conduct quantitative analyses. When participants selected the same articles as the original text, these article choices were coded as 'accurate'. When they provided articles that were not identical to the appropriate articles, these instances were coded as 'inaccurate'. Then, accuracy rates of individual participants were described using descriptive statistics in SPSS. Afterwards, the means and standard deviations of the three proficiency groups were calculated in order to capture the group accuracy scores. Finally, comparisons across groups with regard to the accuracy in the article choices were made by conducting independent samples *t*-tests, an analytical method used to check if the mean scores of two groups on a given variable are

statistically different.

With regard to the second type of data from the think-aloud task, the reasons for article choices reported verbally by the participants were analyzed in the following manner: First, the audio-recorded data obtained from the think-aloud task were transcribed and translated by the researcher. Then, Butler's (2002) taxonomy was employed to code the reasons for article choices into two categories: (1) specific reasons (i.e., the participant was able to identify reasons for using the articles in the text) and (2) non-specific reasons (i.e., the participant could not identify any specific reasons for article choice). Non-specific reasons include article choices based on plausible choices, elimination, intuitive judgments, and guess.

The specific reasons for article choices reported by the participants were further divided into five groups based on the factors the participants reported to have considered when selecting articles for the blanks provided in the text. The description and examples of the five groups are provided in Table 3.7. These reason groups were not established *a priori* but emerged during a preliminary qualitative analysis of the reasons provided by the Korean learners of English during the testing of the think-aloud protocol.

Table 3. 7. Coding for the variable of reasons for article choices in the think-aloud task

Reason category	Description	Examples from the data
Singular vs. plural	An article choice was made based on the grammatical number of a nominal referent.	I chose *a* because it is one banana. (B3) *Mice* is plural, so no article is needed. (I6)
First-mention vs. subsequent mention	An article choice was made based on whether the referent was first introduced or previously mentioned	It needs *the* because the pot is mentioned again. (I1) It is *a* because it is a new noun. (I7)
General referent vs. specific referent	An article choice was made based on the specificity of a nominal referent.	I think it should be *the* because it is a specific cabinet. (B2) He chose a large glass and it's not a specific glass. (NES2)
Countable vs. non-countable	An article choice was made based on the countability of a nominal referent.	You don't need any article because *water* is not countable. (I6)
Combination	An article choice was made based on a combination of two or more of the above reasons.	It is *a* because *lemon* is new and it's singular. (I10)
Non-generalizable or idiosyncratic hypotheses	An article choice was made based on collocation-centered hypotheses or other ad-hoc rules.	I use no article when there is an amount expression. (B9) *The* is used after a preposition. (B7)
Non-specific reason	An article choice was made based on plausible choices, elimination, intuitive judgments, and guess.	Sounds better.. just the way English is used (NES1)

After coding these reasons, descriptive statistical analyses were conducted to understand the distribution of specific and non-specific reasons and rates of accuracy for articles chosen with and without specific reasons. These descriptive statistics provided for an overall description of the kinds of metalinguistic hypotheses generated by each proficiency group.

Next, the frequently cited reasons for the target-like choice of each article form (i.e., *the*, *a*, and $\emptyset$) were examined qualitatively with examples from the think-aloud data. This segment was followed by qualitative analyses of reasons behind non-target-like article choices. These

qualitative analyses were conducted separately, extracting excerpts by the three proficiency

groups in order to determine the kinds of metalinguistic hypotheses that participants generated at

distinctive levels of proficiency. The think-aloud data obtained from the native English speaker

group were analyzed in the same manner as described above for the learner groups. Then, the

reasons provided by the NES participants were compared against those produced by the ESL

participants.

Finally, comparisons between implicit and explicit knowledge of articles were made in a

qualitative manner using three participant cases – one for each proficiency group. First,

metalinguistic hypotheses expressed by the participant in the think-aloud protocol were

determined. Then, the manners in which articles were produced in the oral narrative were

described qualitatively. This segment was followed by comparative discussions on how explicit

knowledge was operationalized in spontaneous oral production.

In sum, this chapter provided a detailed description of the methodological design of this

current study, data collection procedures and data analysis procedures. In the next chapters,

results of the various analyses conducted are reported and discussed to address the five research

questions established for this study.

Chapter IV. Results of Implicit Knowledge of English Articles:
Oral and Written Production of Articles

This Chapter presents the results of the quantitative analyses conducted on the use of

English articles in two kinds of language production data: Oral narratives and written narratives.

Examining articles employed by the participants in speech and writing is an effective way to tap

into the implicit knowledge of English articles possessed by the participants (Akakura, 2009). As

described in Chapter III, oral narratives were obtained by requesting the participants to narrate a story that they thought interesting or meaningful. Written narratives were obtained by having the participants write narratives based on the oral narratives that they produced earlier. The English articles produced were then analyzed with the purpose of addressing research questions that are concerned with the following:

(1) The patterns of article uses according to English language proficiency levels

(2) The differences and similarities in oral and written production of articles.

The other research questions pertaining to metalinguistic hypotheses of articles will be addressed in the next chapter.

The Chapter is organized as follows: The first segment discusses how articles were used in oral narratives in relation to proficiency levels. This section is followed by a general discussion regarding how articles were used in written narratives. The last section draws a comparison between the production of articles in the oral narratives and the written narratives according to proficiency levels.

4.1. Articles Used in Oral Narratives

     4.1.1. Introduction.

This section presents the analyses of English articles used in the corpus of oral narratives with the purpose of addressing the first research question which focused on identifying patterns of article use according to English language proficiency. In analyzing the data of how articles were produced by the participants in each proficiency group, four types of analyses were conducted to examine the following aspects of article use:

(1) the accuracy with which articles were used,

(2) the articles that speakers produced as opposed to the required articles in a given context,

(3) the effect of the appearance of adjectives in noun phrases (NPs) on the accurate

   production of articles, and

(4) the effect of NP contexts on the accurate use of articles.

As discussed in detail in Chapter II, this study adopts the view that the so-called zero article does not exist in English. According to Berezowski (2009), only two overt articles exist in English – the definite article *the* and the indefinite article *a*. In line with Berezowski's theory of the incomplete grammaticalization model, articles only refer to the definite article (*the*) and the indefinite article (*a/an*) in this study. It is acknowledged that the indefinite article has two phonological realizations: *a* before a consonant sound and *an* before a vowel sound. In this study, *a* is used as a cover-term for both phonological forms throughout the Chapter. Likewise, the definite article has two pronunciations: [ðə] before a consonant sound and [ði] before a vowel sound. Since these two phonological variations share the same spelling, *the* is used as a cover-term for both phonological realities. Contexts in which neither the definite article nor the indefinite article is required are referred to as "bare NP" contexts throughout the Chapter.

The organization of Section 4.1 is as follows. First, the data of how English articles were used in the oral narratives by the beginner group (henceforth 'beginners'), the intermediate group ('intermediates'), and the advanced group ('advanced speakers') are analyzed. Second, the patterns that emerged are compared among the three proficiency groups. Finally, key findings are presented.

4.1.2. Beginners: patterns of article use in oral narratives.

This section discusses how beginners used articles in oral narratives. As stated in the introduction section above, the first variable to be examined is accuracy with which the beginners produced the appropriate determiners in NPs in the oral narratives. As discussed in details in Chapter III, *required determiner* is defined as a determiner form that is deemed appropriate in a given NP context. In English syntax, articles are classified as determiners along with other forms such as possessives, demonstratives, and quantifiers. As such, determiners were included in the analysis to examine whether there is a distinct difference in the accurate use of various types of determiners. *N* represents the number of NPs that call for the required determiner. *Accuracy* in this table represents the percentage of instances in which the required determiners were produced appropriately.

Table 4. 1. The accuracy in which beginners produced determiners in oral narratives

Required determiner	*N*	Accuracy (%)
Definite article (*the*)	181	54.1
Indefinite article (*a*)	153	27.5
Possessives (e.g. *my*, *his*)	153	93.5
Demonstratives (e.g. *this*, *that*)	33	90.9
Quantifiers (e.g. *three*, *many*)	32	84.4
Others (e.g. *each, another*)	17	82.4
Bare NP	99	82.8

In Table 4.1, we note that the accuracy for the use of *the* is nearly twice higher than that for *a*. Specifically, the beginners employed *the* with an accuracy of 54.1% and *a* with an accuracy of only 27.5%. The reasons for the higher accuracy in the use of *the* may possibly be related to the different levels of complexity in determining instances in which *the* is required as opposed to *a*. For instance, according to Master's (1990) binary schema of English articles, in

order to produce *the* accurately, speakers need to determine whether a given NP context is definite or not. In order to produce *a* accurately, however, three distinctions need to be made: definiteness, countability, and noun number of NPs. As Trenkic (2008) has maintained, determining all three aspects of NPs at once in spontaneous speech in L2 imposes cognitive challenges to L2 speakers. At the beginner level of proficiency, we see in this table that the participants might have failed to discern contexts that require *a*, resulting in a very low accuracy rate of this form.

While the accuracy for articles was relatively low, the accuracy for non-article determiners was found to be consistently high. The accuracy for possessives was 93.5%, demonstratives, 90.9%, quantifiers, 84.4%, and others, 82.4%. This means that the beginners were able to use these determiners more accurately in spontaneous speech. It is plausible that the frequency in accuracy shown for these determiners is high because their equivalent forms exist in Korean. This observation supports findings of Ekiert (2010), who noted that L2 English learners, whose L1s are Slavic languages that do not employ articles, were able to use non-article determiners more frequently and more accurately.

Finally, in contexts in which no article was required, the rate of accuracy was 82.8%. However, caution is advised in interpreting this frequency because the high accuracy in bare NP contexts may not necessarily suggest that beginners can accurately identify instances in which no article is required. In other words, the beginners may not have produced any article in bare NP contexts for different reasons. In fact, the high frequency in these contexts may be related to a positive L1 transfer that facilitates L2 learning. Ellis (1994) explains that L1 can serve as a facilitative role in L2 acquisition when there is a similarity between L1 and the target language. Bare NP contexts may provide an opportunity for such positive L1 influence. For example, in

Korean, definiteness status of nominals is not marked with overt grammatical markers but is inferred by the hearer or the reader based on textual, contextual, and pragmatic cues. Therefore, like many English learners whose L1s do not have articles, Korean learners of English need to learn that English utilizes articles as overt markers of definiteness and indefiniteness of nominals. However, beginner learners may not be fully aware of the need for articles. Further, L1 influence is stronger at the beginning stages of L2 learning. Thus, they may exhibit a tendency to produce no article. In bare NP contexts in which no article is required, such a tendency of the beginners can result in a positive outcome, as in the case of this beginner group.

The next table draws attention to the distribution of the articles used in relation to the required articles. The variable of *required articles* was cross-tabulated with the variable of *produced determiners*. In SPSS, a cross-tabulation is used to analyze relationships between two variables. It allows a researcher to examine the effect that one variable has on another. In this case, the dependent variable was the accurate production of articles and the independent variable as the required articles. The category of *Others* which appear in *frequency of produced determiner* in the table refers to the use of determiners other than articles (i.e., possessives, demonstratives and quantifiers).

Table 4. 2. The intersection between required articles and produced determiners in beginners' oral narratives

Required article	Frequency of produced determiner				
	the	*a*	omission[4]	others	Total
the	54.1%	1.1%	39.8%	5.0%	100.0% ($N = 181$)
a	12.4%	27.5%	54.9%	5.2%	100.0% ($N = 153$)

$p < .0001$[5]

---

[4] *omission* refers to an instance in which the required article was not produced.

[5] In this study, the *p* value for the variables was found at $p < .0001$, unless otherwise indicated. This value points to statistical significance in the results of the cross-tabulations conducted in this study.

The purpose of Table 4.2 is to illustrate the articles that were required in a context as opposed to those that were supplied by the beginners. In the current study, instances of inappropriate article use were divided into two types: (1) instances of *article omission* when no article is produced in contexts in which the definite article or the indefinite article is required, and (2) instances of *over-generalization* when the definite article or the indefinite article is used in contexts in which they are not required (e.g., using *the* when *a* is required, or using *a* for *the*).

In this table, we find that the most frequently occurring inappropriate article use category in the beginners' oral narratives was article omission.[6] For example, in contexts that required the use of *the*, the beginners produced *the* with an accuracy of 54.1%. Alternatively, when the required article was not employed, the form was omitted with the frequency of 39.8%. A similar pattern was observed in the case of the indefinite article. In contexts that required the use of *a*, the beginners produced *a* with an accuracy of only 27.5%. Instead, they favored omissions in 54.9% of the NPs. As many researchers suggest, the tendency to omit articles could possibly due to an L1 transfer effect since Korean does not have a similar article system (i.e., Ekiert, 2010; Park, 2006; Robertson, 2000; Thomas, 1989; Zegarac, 2004).

With regard to patterns of over-generalization, the definite article was over-generalized more frequently than the indefinite article in the beginners' oral narratives. For example, *a* was produced in only 1.1% of the NPs that required *the*. On the other hand, *the* was produced in 12.4% of the contexts that required *a*. The tendency for L2 English speakers to over-generalize

---

[6] In article research, when no article is produced in contexts in which an overt article is required, some scholars view them as article omission and others, as over-use of the zero article (e.g., Master, 1987; Parrish, 1987; Thomas, 1989). Since the zero article is not assumed in this study, I refer to them as article omission, following a number of scholars (e.g., Butler, 2002; Ekiert, 2010; Jarvis, 2002; Park, 2006; Robertson, 2000; Trenkic, 2007; White, 2003, Young 1996, among others).

the when it is not required has been previously documented by several scholars (Hawkins et al., 2006; Ionin et al., 2004; Parrish, 1987; Zegarac, 2004). In particular, several studies (Ionin et al., 2004; Park, 2006; Thomas, 1989; Zegarac, 2004) found that English learners tend to produce *the* inaccurately when noun referents were indefinite but specific. In the following examples, the underlined nominals were mentioned for the first time in their respective contexts. In these instances, *a* was required for singular nominals and no article was required for plural or uncountable nominals. Instead, beginners employed the definite article.

    (1)   I started to prepare my own script because I saw **the advertisement* in newspaper (B10)
    (2)   That was one part of my job, meeting **the patients*.. (B6)
    (3)   I decided to go* *the big retreat* for mission (B4)

In the above examples, several beginners tended to favor *the* when noun referents were specific in their minds. For instance, in example (3), the retreat that B4 attended was personally very important, so the referent was very clear and specific in her mind, which led her to use *the*. However, the participant did not seem to understand that a specific referent takes *a* when it is not already specific and identifiable in the mind of the hearer.

    The next variable to be examined is the presence of adjectives in NPs. As discussed in Chapter II, Trenkic (2007) argued that L2 learners of English use articles more accurately when nouns are not modified by adjectives. Thus, this independent variable was designed to determine whether adjectives have any effect on the accurate use of the two overt articles in the oral narratives.

Table 4. 3. The distribution of articles produced by beginners in the presence and absence of adjectives in NP

| Required article | Frequency of produced article | | | | | |
| | [+ Adjective in NP] ($N = 101$) | | | [- Adjective in NP] ($N = 561$) | | |
	the	*a*	omission	*the*	*a*	omission
the	36.7%	0.0%	50.0%	57.6%	1.3%	37.7%
a	7.1%	16.1%	73.2%	15.5%	34.0%	44.3%

$p < .0001$

In this table, we find that the beginners produced the definite and indefinite articles more accurately in contexts in which adjectives were not present in the NPs. For instance, *the* was produced with an accuracy of 36.7% in [+ Adjective in NP] contexts, but the accuracy for *the* increased to 57.6% in [- Adjective in NP] contexts. The same pattern can be observed in the production of *a*. In NP contexts that contained adjectives, the accuracy for *a* was only 16.1%. On the other hand, in instances in which *a* was not accompanied with an adjective, the accuracy doubled to 34.0%.

This finding is suggestive of a working hypothesis which posits that the lower accuracy obtained for *the* and *a* in [+ Adjective in NP] contexts is related to the beginners' tendency to omit articles in instances in which a pre-noun position is occupied with a lexical item. In other words, adjectives may compete with articles for the pre-nominal position within NPs. Given limited cognitive resources for language processing, L2 English users may choose to express adjectives instead of articles because adjectives carry more semantic weight. Similarly, Trenkic (2009) explained that the more elements of meaning a speaker needs to encode in a single phrase, the more complex and costly the task becomes. Therefore, she argued that processing constraints are the cause of article omissions in [+ Adjectives in NP] contexts.

In the next table, the effect of various NP contexts on the accuracy of article use is

examined. As discussed in length in Chapter II, research has suggested that the accurate use of articles in English is dependent on syntactic, semantic and discourse contexts (i.e., Huebner, 1979; Master, 1990; Young, 1996). Therefore, it is important to examine whether different NP contexts have an effect on the accurate use of English articles in the oral narratives.

As explained in Chapter III, the variable of NP contexts was coded into four categories. The definite article *the* is used in anaphoric definite contexts (i.e., second and subsequent mentions of a referent that has been previously introduced in the discourse, as in *I bought a book yesterday. The book was on sale.*) and associative definite contexts (i.e., the definiteness of a noun referent is determined by contextual factors such as shared knowledge and post-adjectival modifications, as in *I bought a book yesterday. The cover was red.*). The indefinite article *a* is used in referential indefinite contexts (i.e., first mention of a referent that is specific yet unknown to the hearer as in *I bought a gift for you*), and non-referential indefinite contexts (i.e., a noun referent that does not refer to any specific entity as in *I feel lonely. I need a friend*). For each NP context, the required articles and the produced articles were cross-tabulated to determine the relationship between the variable of NP contexts and the accurate use of articles.

Table 4. 4. The frequency of articles used in the beginners' oral narratives according to NP contexts

Required article	*N*	NP context	Frequency of produced article		
			the	*a*	omission
the	181	Anaphoric definite	45.2%	1.2%	52.4%
		Associative definite	63.0%	1.1%	29.3%
a	153	Referential indefinite	16.1%	29.0%	45.2%
		Non-referential indefinite	9.0%	27.0%	61.8%

p < .0001

In this table, we find that the beginners' accuracy in the oral production of articles is mediated by the variable of NP context. For instance, when *the* was the required article in a

context and the use of this form was related to the NP that was mentioned previously, beginners

produced the form with an accuracy of 45.2%. In turn, when the production of *the* was related to

an associative definite context, accuracy increased to 63.0%.

A partial explanation for the higher accuracy in the use of *the* in associative definite

contexts can be provided with regard to the findings reported above. It is plausible that the

beginners produced *the* more accurately in associative definite contexts because associative

definite contexts include the use of *the* in formulaic sequences[7]. The following are examples of

formulaic sequences that beginners produced in their oral narratives.

(4)   I took the short film <u>for *the* first time</u>. (B10)
(5)   When I saw <u>*the* bible</u>, I afraid of God. (B4)
(6)   I met her <u>in *the* army</u>. (B6)
(7)   So <u>in *the* morning</u>, my cousin just drop over there. (B5)

The underlined NPs are examples of formulaic language that frequently emerged in the corpus.

The frequent occurrence of formulaic sequences in real-time language production can be

explained in terms of the language processing modes. According to Skehan (1996), spontaneous

language production proceeds through two main processing modes: an analytical processing

mode and a holistic processing mode. In the analytical processing mode, speakers apply rules of

language to formulate utterances. In the holistic processing mode, speakers access and retrieve

lexical items of different size as whole units (Schmitt, 2010; Wray, 2000). Even though the

current study did not code for the use of formulaic sequences, we can cautiously suggest that the

beginners seem to have relied on formulaic sequences in producing oral narratives. Indeed, Gor

and Long (2009) maintained that under pressure to produce stories in L2, learners tend to rely on

---

[7] Wray (2000) provides a summary of characteristics of formulaic language. The author uses the term formulaic
sequences to collectively refer to idiomatic forms of expression, including idioms, collocations, and sentence frames.

expressions that they have memorized. This tendency could possibly explain the higher accuracy which appeared for *the* in associative definite contexts than in anaphoric definite contexts.

With regard to the indefinite article, the accuracy seems to be similar between the two contexts: 29.0% in referential indefinite contexts and 27.0% in non-referential contexts. Nonetheless, the distribution of produced articles differed between the two contexts. For instance, *a* was omitted more frequently in non-referential indefinite contexts (61.8%) than in referential indefinite contexts (45.2%). Since non-referential *a* is used for an unspecified referent that is unknown to both the speaker and the hearer and it is not concerned with referentiality for the purpose of cohesive narratives, beginners might not perceive the need for an article. This explanation may partially explain the higher rate of omission of *a* in non-referential indefinite contexts.

On the other hand, when the beginners favored *the* instead of *a*, these inappropriate uses were more frequently attested in referential indefinite contexts (16.1%) than non-referential indefinite contexts (9.0%). As mentioned in Table 4.2, the over-generalization of *the* in referential indefinite contexts in which specific referents are introduced into the discourse is a widely documented phenomenon (Ionin et al., 2004; Park, 2006; Thomas, 1989; Zegarac, 2004). The tendency to over-use *the* in referential indefinite contexts suggests that beginners tended to consider specificity of a referent as a criteria for the use of *the*.

In summary, beginners used English articles in the following patterns:

(1) *The* was used more accurately than *a*.

(2) The most frequent inaccurate article use was omission of *the* and *a*.

(3) Both *the* and *a* were produced more accurately in [- Adjective in NP] contexts.

(4) *The* was produced more accurately in associative definite NP contexts than in anaphoric definite contexts.

(5) *A* was produced more accurately in referential indefinite contexts than non-referential indefinite contexts.

The next section discusses the findings on article use in the oral narratives produced by the participants with intermediate language proficiency.

4.1.3. Intermediates: patterns of article use in oral narratives.

This section examines how the intermediates used articles in oral narratives. Similar to the previous section, this section looks at article use by the intermediate group in the context of the following variables: produced articles as opposed to required articles, presence or absence of adjectives in NPs, and NP context types. Prior to exploring the effect of these variables, we first discuss the distribution in the use of determiners this proficiency level.

Table 4. 5. The accuracy in which intermediates produced determiners in oral narratives

Required determiner	*N*	Accuracy (%)
Definite article (*the*)	166	63.9
Indefinite article (*a*)	204	34.3
Possessives (e.g. *my*, *his*)	154	96.1
Quantifiers (e.g. *three*, *many*)	92	87.0
Demonstratives (e.g. *this*, *that*)	25	84.0
Others (e.g. *each, another*)	17	76.5
Bare NP	200	80.5

According to Table 4.5, similar to the beginner group, the accuracy with which the intermediates produced *the* was considerably higher than the accuracy for *a*. For instance, the accuracy in the use of *the* was 63.9% and for *a*, 34.3%. We note that the accuracy for *a* is still

relatively low, considering that the speakers in the intermediate group have taken a minimum 6

years of EFL education in Korea and have lived in the United States for a period ranging from

two to five years. As discussed in the beginner section, the complexity of identifying contexts

and accurately producing *a* seems to continue to pose challenges to the intermediates in

spontaneous speech.

With regard to the other kinds of determiners, the data suggest a relatively high accuracy

(i.e., possessives 96.1%, quantifiers 87.0%, demonstratives 84.0%, and others 76.5%). These

patterns suggest that the intermediates have a similar tendency as the beginners with regard to

the use of other kinds of determiners. As suggested earlier, this tendency may be influenced by

the fact that equivalent forms of these determiners exist in Korean.

Finally, the accuracy found in bare NP contexts was very high (80.5%). Similar to the

beginner group, this high accuracy of bare NP contexts in which no article is required may also

be related to positive L1 transfer since their L1 does not utilize articles to mark definiteness of

nominals.

Continuing with a similar analysis to the one conducted in the beginners' section, the

next table illustrates the articles (*a* and *the*) that the intermediates produced in their oral

narratives with respect to the required articles. *Others* include determiners other than articles (i.e.,

possessives, demonstratives, and quantifiers).

Table 4. 6. The intersection between required articles and produced determiners in intermediates'
oral narratives

Required article	Frequency of produced determiner				
	the	*a*	omission	others	total
the	63.9%	0.0%	34.3%	1.8%	100.0% (N = 166)
a	8.3%	34.3%	53.4%	4.0%	100.0% (N = 204)

p < .0001

Table 4.6 illustrates the frequencies of accurate article use as well as inaccurate article use (i.e., instances of article omission and over-generalization). In the intermediate group, the most frequent pattern of non-target-like article use was article omission. For example, in contexts that required *the*, the intermediates were able to produce the required form accurately in 63.9% of the NPs. However, when *the* was not produced, they omitted the definite article in 34.3% of the NPs. Similarly, in NP contexts that required *a*, the form was omitted in a little over a half of the times it was required (53.4%). In short, similar to the beginner group, the tendency to omit articles was evidenced in contexts that required *the* and *a* in the intermediate group, but the tendency was more salient in instances in which *a* was required.

With regard to patterns of article over-generalization, the frequency in which *the* was over-generalized was 8.3% in contexts in which *a* was required. Scholars have noted that as English language learners start to use *the* more frequently, they go through a period of over-generalization of the definite article (e.g., Hawkins et al., 2006; Liu & Gleason, 2002; Zegarac, 2004). Indeed, beginners over-used *the* in 12.4% of the instances in which *a* is required. In contrast, the intermediates were found to make fewer errors of over-generalizing *the* (8.3%). Furthermore, we found no evidence in this group when they produced *a* instead of *the*. This finding suggests that at this level of proficiency, speakers were able to detect the distinction between these two forms.

The next table examines the effect of the variable of the presence of adjectives in NPs on accurate use of articles.

Table 4. 7. The distribution of articles produced by intermediates in the presence and absence of adjectives in NP

| Required article | Frequency of produced article | | | | | |
| | [+ Adjective in NP] ($N = 154$) | | | [- Adjective in NP] ($N = 704$) | | |
	the	*a*	omission	*the*	*a*	omission
the	58.1%	0.0%	41.9%	65.2%	0.0%	32.6%
a	5.0%	18.3%	76.7%	9.7%	41.0%	43.8%

p < .0001

In Table 4.7, we find that the intermediates' ability to produce articles accurately was negatively influenced by the presence of adjectives in NPs. For example, *the* was produced accurately in 58.1% of the [+ Adjective in NP] contexts, but the accuracy increased to 65.2% in [- Adjective in NP] contexts. A similar pattern was observed in the use of *a*. In NP contexts that contained adjectives, the accuracy for *a* was only 18.3% while the accuracy increased to 41.0% in [- Adjective in NP] contexts. As pointed out in the beginner section, the lower accuracy in the production of *the* in [+ Adjective in NP] context can be related to the tendency to omit articles more frequently in NPs that contain adjectives because adjectives seem to compete with articles in a pre-nominal position.

To briefly summarize, the working hypothesis proposed in the beginner section was also supported in the corpus of the oral narratives by the intermediates. That is, the intermediate speakers preferred to omit articles in instances where adjectives were expressed. Further, the negative effect of the presence of adjectives on article use was stronger for the indefinite article *a*. That is, the intermediates tended to omit *a* more readily than *the* possibly because, as Akakura (2009) points out, *a* is considered phonologically and semantically less salient than *the*.

Next, the effect that various NP contexts have on the accurate use of articles is discussed. As described in Chapter III, NP contexts in which articles are required were coded into four

types: anaphoric definite contexts (i.e., second and subsequent mentions of an already known referent) and associative definite contexts (i.e., the definiteness of a noun referent is determined by contextual factors such as shared knowledge and post-adjectival modifications) for the use of *the*, and referential indefinite contexts (i.e., first mention of a referent that is specific yet unknown to the hearer) and non-referential indefinite contexts (i.e., a noun referent that does not refer to any specific entity) for the use of *a*. For each NP context, the distribution of articles produced by the intermediates in their oral narratives is examined in the next table.

Table 4. 8. The frequency of articles in the intermediates' oral narratives according to NP contexts

Required article	N	NP context	Frequency of produced article		
			the	*a*	omission
the	166	Anaphoric definite	58.5%	0.0%	41.5%
		Associative definite	67.0%	0.0%	30.4%
a	204	Referential indefinite	14.5%	31.3%	48.2%
		Non-referential indefinite	4.2%	35.8%	57.5%

p < .0001

According to Table 4.8, the context in which an NP was produced seems to influence the accuracy with which the articles were produced by intermediates. Specifically, when *the* was the required article in a context and the use of this form was related to an NP that was mentioned previously, intermediates produced the form with an accuracy of 58.5%. In turn, when the use of *the* was related to an associative definite context, accuracy was higher (67.0%). As mentioned in the beginner section, this pattern may possibly be attributed to the use of *the* in formulaic sequences. In spontaneous speech, formulaic expressions seem to be produced more frequently in order to improve fluency in speaking. When *the* was part of memorized sequences, the form could likely be used accurately by intermediates in associative definite contexts.

With regard to the indefinite article, the intermediates produced non-referential *a* more

accurately (35.8%) than referential *a* (31.3%). In instances in which *a* was not produced

accurately, *a* was omitted more frequently in non-referential indefinite contexts (57.5%) than in

referential contexts (48.2%). Since non-referential *a* is used for a referent that is not specified

and is unknown to both the speaker and the hearer, intermediates might not have perceived the

need for an article and thus omitted it with a higher frequency. Meanwhile, *a* was substituted

with *the* more frequently in referential contexts (14.5%) than in non-referential contexts (4.2%).

Similar to the beginners, the intermediates tended to perceive referential indefiniteness as

definite contexts due to the specificity or semantic saliency of a noun referent and thus they

favored the use of *the*.

In summary, the intermediate speakers used English articles in the oral narratives in the

following patterns:

(1) *The* was used more accurately than *a*.

(2) When the articles were not produced accurately, the tendency to omit them was

evidenced.

(3) The presence of adjectives had negative effect on the accuracy with which articles

were employed.

(4) *The* was produced more accurately in associative definite contexts than in anaphoric

definite contexts.

(5) *A* was produced more accurately in non-referential indefinite contexts.

In the next section, findings on patterns of article uses in the oral narratives produced by

the advanced participants are reported.

4.1.4. Advanced group: patterns of article use in oral narratives.

This section presents the analyses of the English articles used by the advanced

participants in the corpus of oral narratives. Table 4.9 presents the distribution in the use of

determiners.

Table 4. 9. The accuracy in which the advanced produced determiners in oral narratives

Required determiner	*N*	Accuracy (%)
Definite article (*the*)	252	76.2
Indefinite article (*a*)	205	55.1
Demonstratives (e.g. *this*, *that*)	37	97.3
Possessives (e.g. *my*, *his*)	178	95.5
Quantifiers (e.g. *three*, *many*)	76	93.4
Others (e.g. *each*, *another*)	15	86.7
Bare NP	162	72.8

Table 4.9 shows that *the* was produced with a higher rate of accuracy than *a*. For

instance, the advanced speakers employed *the* in 76.2% of the NPs while they used *a* accurately

in 55.1% of the NPs in the oral narratives, a 20 percentage point difference. A variety of

variables could have contributed to the lower accuracy in the oral production of *a*, including the

presence of adjectives in NPs as well as NP contexts for the indefinite article. In the remaining

parts of this section, these variables are examined to elucidate why the advanced group did not

produce *a* as accurately as *the*.

We also note in Table 4.9 that the accuracy for other categories of determiners was

higher than the accuracy for articles, a finding that was revealed in the previous segments

discussed (i.e., beginners and intermediates). Specifically, demonstratives were used with an

accuracy of 97.3%, possessives, 95.5%, quantifiers, 93.4%, and others, 86.7%. Similar to the

patterns evidenced in the beginners and intermediates, with more certainty, we may attribute the

high rates of accuracy to positive influence of L1, although causal effects do not fall within the scope of this study. We thus can claim that the consistent findings uncovered by all three proficiency levels seem to point to positive L1 transfer with regard to non-article demonstratives.

With respect to bare NP contexts, contexts which do not require an article, the accuracy was found at 72.8%, a slightly lower frequency than the frequencies represented by the beginner and the intermediate groups. As mentioned in the previous sections, the high rates of accuracy in bare NP contexts observed in the lower proficiency groups may be partly attributed to positive L1 transfer. With regard to participants in the advanced group, however, we find that they are more immersed in the target language environment and less under direct L1 influence. They have more target-like use of articles in the oral narratives compared to the other groups and this tendency is evident in bare NP contexts as well.

The next table draws our attention to the use of the two articles in the oral narratives by the advanced speakers. The distribution of the articles used is presented according to the required article and produced determiners.

Table 4. 10. The intersection between required articles and produced determiners in advanced speakers' oral narratives

Required article	Frequency of produced determiner				
	the	*a*	omission	others	total
the	76.2%	0.0%	20.6%	3.2%	100.0% (N = 252)
a	9.8%	55.1%	33.2%	1.9%	100.0% (N = 205)

p < .0001

According to this table, advanced speakers produced *the* in a majority of instances in which the article was required. The frequency of target-like production of *the* was 76.2%, a high rate of accuracy, which suggests that the implicit knowledge of the definite article has reached a stable level in the advanced group. When *the* was not produced as required, it was omitted

(20.6%) rather than replaced by *a*. The fact that they did not substitute *the* with *a* points to the

fact that the advanced speakers are aware of the distinction between these two forms.

With regard to *a*, the accuracy was considerably lower than what we found with respect

to the accuracy for *the*. The target-like production of *a* occurred only in 55.1% of the time, a low

rate of accuracy considering the advanced level of English proficiency in this group. In instances

in which advanced speakers did not use the required article, they tended to omit it (33.2%) more

frequently than over-using *the* (9.8%). Several researchers have documented that *the* tends to be

acquired before *a* (e.g., Huebner, 1983; Parrish, 1987; Thomas, 1989). These observations

suggest that different patterns of production between the two articles seem to lend support to this

suggested order of article acquisition.

The following table shows whether the variable of the presence of adjectives in NP has

an effect on the accuracy of article use.

Table 4. 11. The distribution of articles produced by advanced speakers in the presence and absence of adjectives in the NP

Required article	Frequency of produced article					
	[+ Adjective in NP] (*N* = 180)			[- Adjective in NP] (*N* = 745)		
	the	*a*	omission	*the*	*a*	omission
the	60.6%	0.0%	36.4%	81.7%	0.0%	15.1%
a	5.1%	35.6%	59.3%	11.6%	63.0%	22.6%

p < .0001

Similar to what was revealed among beginners and intermediates, the accurate use of

articles was affected negatively by the presence of adjectives in NPs. For instance, the accuracy

in the production of *the* was 60.6% in [+ Adjective in NP] contexts but increased to 81.7% in [-

Adjective in NP] contexts. Similarly, the accuracy for *a* was 35.6% in [+ Adjective in NP]

contexts but that frequency increased by almost 30 percentage points to 63.0% in [- Adjective in

NP] contexts. These patterns suggest that the presence of adjectives in NPs affects the accuracy of article use by speakers even at the advanced level of proficiency.

Next, the effect of various NP contexts on the accurate use of articles is discussed. As mentioned earlier, various NP contexts were coded into four types: anaphoric definite contexts and associative definite contexts for the definite article, and referential indefinite contexts and non-referential indefinite contexts for the indefinite article. For each NP context, required articles were cross-tabulated with produced articles.

Table 4. 12. The frequency of articles used in the advanced speakers' oral narratives according to NP contexts

Required article	*N*	NP context	Frequency of produced article		
			the	*a*	omission
the	252	Anaphoric definite	84.1%	0.0%	12.7%
		Associative definite	73.6%	0.0%	23.6%
a	205	Referential indefinite	16.3%	55.8%	25.6%
		Non-referential indefinite	4.3%	55.6%	38.5%

p < .0001

A general finding we revealed in this table points to the fact once again that the accuracy in the production of articles varied according to different NP contexts. For instance, with regard to the definite article, the production of the definite article *the* was found more accurate in anaphoric definite contexts (84.1%) than in associative definite contexts (73.6%), both relatively high frequencies. This observation suggests that the advanced speakers are able to discern *the* accurately when a noun referent has been already introduced and is subsequently mentioned, a tendency not revealed in the beginners and intermediates. More specifically, they reached a threshold in which they are able to make a distinction between instances in which an NP was previously mentioned and instances in which an NP appeared in associative definite contexts.

With regard to the NP contexts for the indefinite article, the advanced speakers used *a*

with similar rates of accuracy in the two different contexts (55.8% in referential indefinite contexts and 55.6% in non-referential indefinite contexts). However, different patterns of inaccurate article use emerged for each indefinite context. That is, there were more instances of over-generalization of *the* in referential indefinite contexts while more frequent omissions of *a* occurred in non-referential indefinite contexts. As evidenced previously with beginners and intermediates, the advanced speakers over-generalized *the* more frequently in referential contexts (16.3%) than in non-referential contexts (4.3%). Referential indefinite contexts refer to instances in which specific referents are introduced for the first time in the discourse. Due to the nature of specificity inherent in referents, the advanced participants may have produced *the* inaccurately by making associations between the definite article and specificity of referent.

Different patterns of article use were revealed in non-referential indefinite contexts. The advanced participants omitted *a* more frequently in non-referential contexts (38.5%) than in referential contexts (25.6%). Since we understand that non-referential use of *a* is not critical for referential coherence in a discourse, the speakers may have omitted *a* more readily in spontaneous language production. Again, these findings call into question Ekiert's (2010) study. Ekiert found that advanced L2 English speakers of Slavic languages omitted *a* more frequently in referential contexts. In this current study, *a* was omitted more frequently in non-referential contexts. The possible reasons for these tendencies will be discussed in the section on qualitative data from the think-aloud protocol which appears in Chapter V. In the think-aloud task, we elucidate further on why they may have exhibited this tendency.

In summary, the advanced group used English articles in the oral narratives in the following patterns:

(1) *The* was produced with the highest rate of accuracy of the two forms.

(2) When *the* and *a* were not used accurately, a general favoring to omit articles was evidenced.

(3) The presence of adjectives in NPs negatively influenced accuracy of article use.

(4) *The* was produced more accurately in contexts in which the referent had been previously mentioned.

(5) *A* was produce with similar rates of accuracy in referential and non-referential indefinite contexts, but the form was omitted more frequently in non-referential contexts.

In the next section, the patterns in which the participants at all levels of proficiency used English articles are compared among the three proficiency groups. The comparative analyses are conducted to address the first research question that focuses on determining patterns of article use along the proficiency continuum.

4.1.5. Group comparisons of article use in oral narratives.

In this section, a comparison is made between the three groups of participants with regard to how they used articles in their oral narratives. The following figure illustrates the accuracy patterns for two articles (*the*, *a*) according to the external variable of the participants' proficiency.

Figure 4. 1. The article accuracy according to proficiency
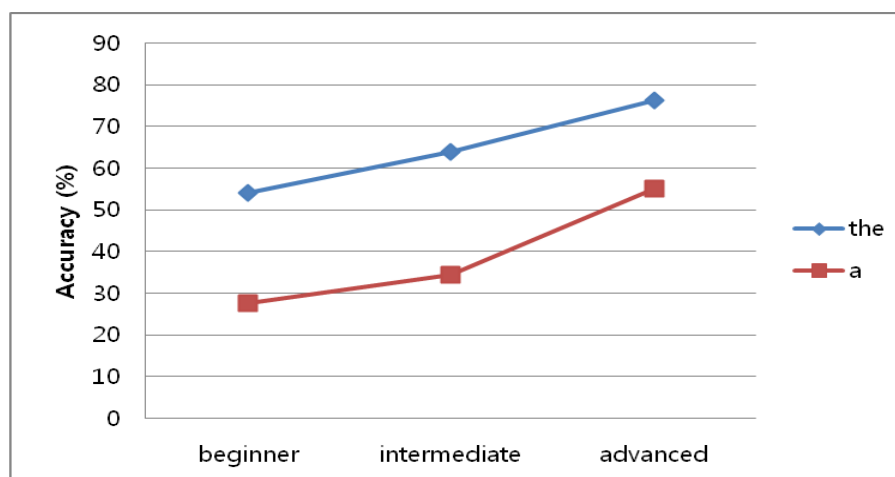
At a glance, and as was expected, we find that accuracy for *the* and *a* increases along the proficiency continuum, a finding supported by previous research (Huebner, 1979; Oller & Redding, 1971; Park, 2006; Parrish, 1987; Young, 1996). The figure also illustrates that all proficiency groups exhibited less accuracy in their production of the indefinite article *a* in the oral narratives. The difference in accuracy between the two articles seems to persist even at the advanced level of proficiency. The marked gap in accuracy between the two articles may be related to the manner in which definiteness and indefiniteness are marked in Korean. According to Kim (1990), definiteness is indicated by demonstratives such as *–geu* in some cases, but there is no morpheme equivalent of *a* that indicates indefiniteness. In other words, Korean speakers of English may associate *the* with the use of certain demonstratives in Korean, but they do not have a form in their L1 to which they may map onto for the use of *a*. This may be one reason why *a* is more difficult for participants to orally produce accurately.

Tables 4.13 and 4.14 draw our attention to the frequencies of non-target-like article use in the three proficiency groups. Table 4.13 shows the frequencies of inaccurate article production in contexts in which *the* was required. We illustrate this table with the purpose of showing how the

three groups tended to perform in a similar manner with regard to omitting an article when the environment required *the*.

Table 4. 13. The frequency of non-target-like article production in contexts that required *the*

	Accuracy of *the*	Omission of *the*	Over-generalization of *a*
Beginner	54.1%	39.8%	1.1%
Intermediate	63.9%	34.3%	0.0%
Advanced	76.2%	20.6%	0.0%

According to this table, the accuracy with which *the* was produced improved according to proficiency levels. Specifically, beginners employed *the* with 54.1% of accuracy, intermediates with 63.9% of accuracy, and advanced speakers with 76.2% of accuracy. Similarly, the frequency of omitting *the* decreased along the proficiency continuum. Beginners omitted *the* in 39.8% of the cases, intermediates, in 34.3%, and advanced speakers, much less in 20.6%. Finally, frequency of over-generalizing *a* in instances in which *the* was required was very low.

Table 4. 14. The frequency of non-target-like article production in contexts that required *a*

	Accuracy of *a*	Omission of *a*	Over-generalization of *the*
Beginner	27.5%	54.9%	12.4%
Intermediate	34.3%	53.4%	8.3%
Advanced	55.1%	33.2%	9.8%

Several observations can be made from this table. Firstly, there was a positive correlation between accuracy in the use of *a* and the external variable of proficiency levels. For intance, beginners produced *a* in only 27.5% of the instances, but intermediates, in 34.3% of the instances, and advanced speakers, higher in 55.1% of the instances. Secondly, the beginners and intermediates omitted *a* more frequently (54.9% and 53.4%, respectively) than producting it. As noted by Akakura (2009), the high frequency of the omission of *a* may be attributed to the concerns of saliency. That is, the indefinite article makes less significant semantic contribution to

the construction of the overall message than the definite article. Further, *a* is often phonologically reduced in natural speech. As such, L2 English speakers whose L1s do not have an article system may tend to omit it more readily in carrying out a cognitively challenging oral narrative task. However, this tendency was reversed in the advanced group since the target-like production of *a* was more frequent (55.1%) than its omission (33.2%).

Finally, we observe the fluctuations in the frequencies with respect to the over-generalization of *the* in contexts in which *a* was required. For example, the beginners over-generalized *the* in 12.4% of the NPs, and the frequency decreased to 8.3% in the intermediate group. However, it increased again to 9.8% in the advanced group. The slightly higher frequence which which the advanced group over-generalized *the* may be related to their preference for the use of *the* in instances in which specific referents are mentioned for the first time, a tendency which will be discussed later in Table 4.15.

In the next figure, the effect of the presence of adjectives in the NPs on the patterns of article use in the oral narratives for all three groups is illustrated.

Figure 4. 2. The accuracy in article use in the presence and absence of adjectives in the NPs
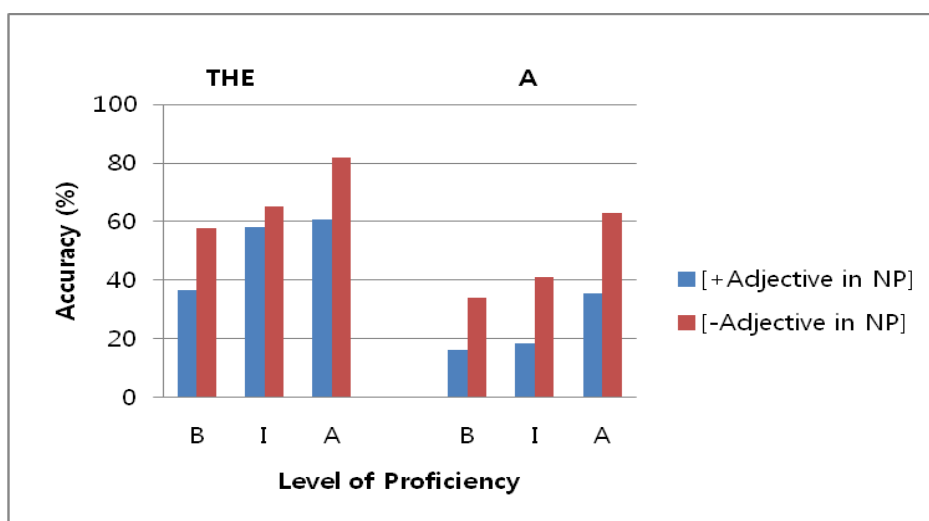
In this figure, we find that the accuracy in the use of *the* and *a* was higher in [- Adjective in NP] contexts and that this observation was evident at all levels of proficiency. With regard to this observation of negative influence of the presence of adjectives in NPs, a working hypothesis was proposed in the section on the beginner group. It was hypothesized that when both an adjective and an article compete in the pre-nominal position, L2 English learners tend to favor the production of the adjective rather than the article. It is plausible that this general favoring occurs because second language learners have limited capacity for language processing in spontaneous language production. For instance, Trenkic (2009) made a similar claim that whenever cognitive resources are exceeded, as in the more demanding sequence of article + adjective + noun, articles fail to be produced, resulting in higher instances of article omission. A second explanation is also plausible. Given that the two competing forms are distinct, one more semantically salient than the other, L2 Korean speakers opt to use the more semantically salient form. In other words, adjectives override articles. Furthermore, adjectives are a syntactic category that exists in Korean, while articles are not. Therefore, given the above reasons, Korean learners of English may opt for an adjective over an article in spontaneous oral production when both forms emerge in the pre-posed position.

The negative effect of the presence of adjectival modifications in NPs was stronger on the accurate use of *a* as measured in frequencies. As shown in the figure, the difference in accuracy for *a* between [+ Adjective in NP] contexts and [- Adjective in NP] contexts is found to be higher in all proficiency groups. A couple of points can be made regarding the observed pattern. The first point is related to the observed higher rates of omission of *a* when adjectives were present in NPs. In this case, the participants might have omitted *a* more frequently than *the* when an adjective competed syntactically for the pre-nominal position because *a* is semantically and

phonologically less salient than *the*. The second point is concerned with the higher rates of overuse of *the* instead of *a* when adjectives were used in NPs. In other words, participants tended to over-generalize *the* in instances in which *a* was required, and this tendency was more frequent in [+ Adjective in NP] contexts. In this case, participants may have considered that the adjectival modifications make noun referent more specific, thus requiring the use of *the*. In other words, adjectives provide a more specific description of noun referents, and participants might conclude that these referents were definite and thus, a more specific form was required. These two plausible explanations will be discussed in more details with the qualitative data from the think-aloud protocol in Chapter V when we ask the speakers to detail the reasons why they chose one form over the other.

In the next segment, I discuss how the variable of NP contexts affects accuracy. Figure 4.3 compares accuracy for *the* in two NP contexts among the three proficiency levels.

Figure 4. 3. The accuracy for *the* according to NP contexts
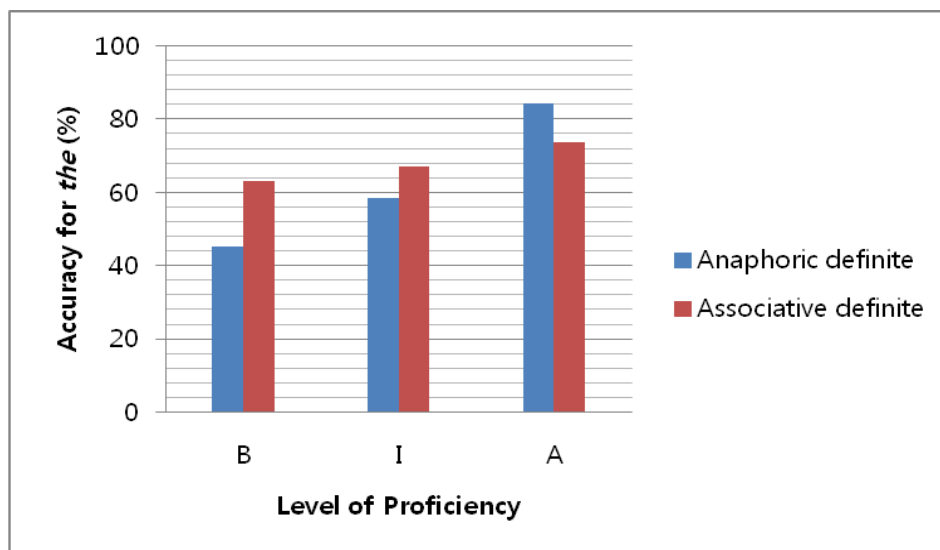
Figure 4.3 is illustrative of the tendencies that emerged in the tables, namely, that *the* is used with more accuracy by beginners and intermediates in NPs when there is an associative

definiteness (i.e., definiteness of a noun referent is determined by various contextual factors such as common knowledge, physical presence of a referent, and post-nominal modifications). The following are examples of instances from the corpus in which *the* was used in associative definite contexts.

(8)   Suddenly, water busted from there (…) *the floor* was wet (B9) (definite by way of association)

(9)   This night is *the last chance* to say anything to her (I3) (assumed to be unique with the superlative)

(10)   The sand is really really hot (…) it's funny *the surface* is really hot but *the inside* is kind of cold (I4) (definite by way of association)

As mentioned in the individual proficiency group sections, many of the associative definite contexts in which *the* was used accurately were related to the use of *the* in idioms and formulaic expressions. As such, even the beginners were able to produce *the*, to some degree, accurately in these memorized phrases. Therefore, a gradual increase in the accuracy for *the* can be observed in associative definite contexts in this figure.

While the increase in the accuracy for *the* in associative definite contexts was gradual, the increase was more dramatic in anaphoric definite contexts. The accurate use of *the* in anaphoric definite contexts requires the speaker to determine if a noun referent was already introduced in the context and how far back was the referent or the form of a noun mentioned to determine referential coherence. In this figure, it is evident that the beginners still lack the ability to produce *the* accurately when a referent has been previously mentioned. Although beginners seem to exhibit some evidence that signals low levels of accuracy, they have not reached a threshold that points to a firmly situated acquisition of this form in anaphoric definite contexts. By comparison, *the* was produced by the advanced participants with a high rate of accuracy in

anaphoric definite contexts. This pattern suggests that the accurate use of *the* in contexts of

second mentions of previously introduced referents is a stable part of the implicit knowledge of

English assumed by the advanced participants.

The next figure illustrates accuracy for *a* in two NP contexts: referential indefinite

contexts (i.e., the use of *a* to introduce a new specific entity into the discourse) and non-

referential indefinite contexts (i.e., attributive and predicative uses of *a* when the noun phrase

does not refer to any specific entity). Unlike the NP contexts for *the*, the comparison of the effect

of NP contexts on the use of *a* is presented in a table format to illustrate different frequencies of

inaccurate article use depending on NP contexts for all three groups.

Table 4. 15. Frequencies of accurate use of *a* and inaccurate article use in two NP contexts that require the indefinite article

NP context	Proficiency	Accuracy	Omission of *a*	Over-generalization of *the*
Referential indefinite context	Beginner	29.0%	45.2%	16.1%
	Intermediate	31.3%	48.2%	14.5%
	Advanced	55.8%	25.6%	16.3%
Non-referential indefinite context	Beginner	27.0%	61.8%	9.0%
	Intermediate	35.8%	57.5%	4.2%
	Advanced	55.6%	38.5%	4.3%

In this table, we first note that, at all proficiency levels, omissions were more frequent in

non-referential indefinite contexts than referential indefinite contexts (61.8% vs. 45.2% in the

beginner group, 57.5% vs. 48.2% in the intermediate group, and 38.5% vs. 25.6% in the

advanced group). Consider the following examples from the corpus of oral narratives produced

by advanced speakers. The symbol ∅ is used to indicate that an article is missing.

(11)   My mother was *∅ English teacher (A9)
(12)   It was *∅ really precious moment (A2)
(13)   It was *∅ very tight schedule (A4)

134

Similar to the above examples, many instances in which *a* was omitted contained adjectives because non-referential indefinite contexts include attributive uses as a way to describe a referent (e.g., *This is a beautiful dress*). As illustrated in Figure 4.2, the presence of adjectives in NPs had more negative effect on the accurate use of *a*. Therefore, the more frequent omissions of *a* in non-referential contexts may be related to the effect of the presence of adjectives in NPs on the use of *a*. Nonetheless, it should be noted that a correlation exists: the frequency of omitting *a* decreased with increased proficiency. That is, the higher the speaker was on the scale of proficiency, the more likely they were able to identify contexts that require *a* and produce the form accurately.

Alternatively, instances of the over-generalization of *the* were more frequent in referential indefinite contexts in all proficiency groups (16.1% vs. 9.0% in the beginner group, 14.5% vs. 4.2% in the intermediate group, and 16.3% vs. 4.3% in the advanced group). In other words, participants mistakenly judged referential indefinite contexts as definite contexts more often than they did with non-referential indefinite contexts. As mentioned in the previous sections that presented different group tendencies, this finding is supported by what has been documented by scholars: the tendency to over-generalize *the* when noun referents are specific (e.g., Hawkins et al., 2006; Ionin et al., 2004; Liu & Gleason, 2002; Parrish, 1987; Thomas, 1989; Zegarac, 2004). The following are examples in which participants produced *the* instead of *a* in contexts in which specific referents were introduced for the first time in the discourse.

(14) I enrolled in **the flight school* in Princeton and I kind of liked it. (A6)
(15) The owner caught **the taxi* for me and we went together. (I9)
(16) We went to **the bar*.. famous bar in Manhattan. (B2)
(17) I would like to talk about **the woman* I met. (B6)
(18) I saw **the advertisement* in newspaper (B10)

In these examples, *a* was needed because the referents were introduced for the first time in the discourse, but *the* was produced instead because these participants possibly made a mistake of judging them as definite nominals. In English, a referent is considered indefinite when it is not previously known to the hearer, even though the referent is specific in the mind of the speaker. However, L2 learners of English tend to associate the definite article with specificity of the referent rather than the hearer's familiarity with the referent.

    4.1.6. Summary of key findings: article use in oral narratives.

    In this section, patterns of article use in the oral narratives were examined in three proficiency groups for the purpose of addressing the first research question that focused on identifying how articles are produced at different levels of proficiency. Specifically, articles produced by the participants were analyzed with regard to the variables of overall accuracy, patterns of non-target-like article use, the presence of adjectives in NPs, and the categories of NP contexts. Key findings can be summarized as the following:

(1) There was a positive correlation between accuracy in article use and proficiency level.

(2) Article omissions were the most frequently non-target-like article use for both article forms (*the* and *a*) in the oral discourse.

(3) In all proficiency groups, *the* was produced more accurately than *a*.

(4) In all proficiency groups, articles were omitted more frequently when adjectives were present in NPs than when no adjectives were used in NPs.

(5) The tendency to omit articles in NPs with adjectives was more evident with the indefinite article than the definite article.

(6) With regard to NP contexts for the definite article, beginners and intermediate

participants produced *the* more accurately in associative definite contexts than in anaphoric definite contexts. Conversely, advanced speakers employed *the* more accurately in anaphoric definite contexts in the oral narratives.

(7) With regard to NP contexts for the indefinite article, rates of accuracy were similar between referential and non-referential indefinite contexts.

(8) In all groups, *a* was omitted more frequently in non-referential indefinite contexts.

(9) In all groups, *a* was substituted with *the* more frequently in referential contexts.

The patterns of article that emerged in the analysis of the oral corpus were summarized in the above. The extensive discussion of these findings in the context of the previous research is provided in Chapter VI. Now we turn to patterns of article use in the written narratives in the next section.

4.2. Articles Used in Written Narratives

4.2.1. Introduction.

This section presents the analyses of English articles produced in the corpus of written narratives. It addresses the first research question which focused on determining the patterns of article use in oral and written discourses according to English language proficiency. In the data collection process, oral narratives were first obtained by requesting the participants to narrate a story that they thought interesting or meaningful to them. Then, the participants wrote narratives based on the oral narratives that they produced earlier. Time limit was not placed on the writing task to allow the participants sufficient time to monitor their written production.

In this section, similar to the previous section on oral narratives, four analyses were conducted to examine the following aspects of article use:

(1) the accuracy with which articles were used,

(2) the articles produced as opposed to the required articles in a given context,

(3) the effect of the appearance of adjectives in NPs on the accurate production of articles, and

(4) the effect that NP contexts had on the accurate use of articles.

Similar to the previous section on article use in oral narratives, this section is informed by Berezowski's (2009) perspective that the English article system consists of the definite article (*the*) and the indefinite article (*a/an*). Thus, findings and discussions presented in this section pertain only to the two overt articles without making reference to the zero article. In English, there are a variety of linguistic environments in which articles are not allowed, which Berezowski (2009) claims to be the result of the on-going grammaticalization process of the two articles. These contexts in which articles are inadmissible are referred to as "bare NP" contexts in the present study.

The organization of Section 4.2 is as follows. First, the data of how English articles were used in the written narratives by the beginners, intermediates, and advanced participants are analyzed. Second, the patterns that emerged among the three groups are compared. Finally, key findings are discussed.

4.2.2. Beginners: patterns of article use in the written narratives.

This section discusses how beginners used articles in written narratives. As stated in the introduction above, the first variable to be examined is the accuracy with which determiners were produced in the beginners' written narratives.

Table 4. 16. The accuracy in which beginners produced determiners in written narratives

Required determiner	*N*	Accuracy (%)
Definite article (*the*)	152	55.9
Indefinite article (*a*)	140	38.6
Demonstratives (e.g. *this*, *that*)	21	100.0
Quantifiers (e.g. *three*, *many*)	16	100.0
Others (e.g. *each*, *another*)	12	100.0
Possessives (e.g. *my*, *his*)	124	92.7
Bare NP	88	86.4

From this table, we find that the accuracy in the use of *the* is considerably higher than that for *a*. The beginners employed *the* with an accuracy of 55.9% and *a* with an accuracy of only 38.6% in their writing. The pattern of higher accuracy in the use of *the* was also evident in the beginners' oral narratives. In fact, this pattern may support what has been documented related to the order of article acquisition. That is, several scholars have reported that in general, *the* is acquired before *a* (Master, 1987; Park, 2006; Thomas, 1989, among others). One reason for the delayed acquisition of the indefinite article is that the accurate production of *a* requires the speaker to make distinctions on definiteness, countability, and noun number of NPs. When beginners did not have sufficient understanding of how *a* was used in English, they were probably not able to accurately discern contexts that required *a* even though they were afforded sufficient time to reflection on their written language for grammatical accuracy.

While the frequency of the accurate use of the two articles was relatively low, the accuracy for other types of determiners was found to be consistently high. For instance, the accuracy for demonstratives, quantifiers and others was 100%. Possessives were produced with an accuracy of 92.7%. Thus, with these data and the fact that similarities emerged with regard to these forms in oral narratives, we can now suggest that the accuracy in using these determiners is

high because their equivalent forms exist in Korean. In a similar vein, Ekiert (2010) also found that non-article determiners were produced fairly accurately by the participants from article-less Slavic L1 backgrounds. Thus, the facilitative effect of the L1 seems to mediate the mapping between forms and meanings of determiners in English at the beginner level of proficiency.

Finally, the accuracy in bare NP contexts is much higher than the accuracy for the overt articles. For example, the beginners did not produce any article in 86.4% of the bare NP contexts. However, this does not necessarily suggest that beginners are able to identify contexts in which no article is needed. Rather, the high frequency with which beginners produced no article in bare NP contexts may be attributed to the influence of their L1s that do not have articles. That is, these participants are at the beginning stage of the second language acquisition, in which L1 influence is relatively strong. As a result, beginners may have produced no article simply because of L1 effect. It should be noted that these instances in which no article was employed were coded as "accurate" use in bare NP contexts. However, it is not clear whether the accuracy was mediated by the participants' knowledge of English or as a consequence of positive L1 influence.

As the above table illustrated low rates of accuracy in the written production of articles in the beginner group, it is important to examine what forms beginners employed when they did not produce the required articles. The following table illustrates the distribution of produced determiners in relation to the required articles. In this table, the category of *Others* refers to the use of determiners other than articles (i.e., possessives, quantifiers and demonstratives).

Table 4. 17. The intersection between required articles and produced determiners in beginners' written narratives

Required article	Frequency of produced determiner				
	the	*a*	omission	others	Total
the	55.9%	1.3%	37.5%	5.3%	100.0% ($N = 152$)
a	7.9%	38.6%	48.6%	4.9%	100.0% ($N = 140$)

$p < .0001$

Table 4.17 shows patterns of accurate articles in written production as well as inaccurate article production when the required articles were not used. As discussed in the section on oral narratives, inaccurate article use was divided into two types: article omission (i.e., the required article is omitted), and article over-generalization (i.e. *the* is used inaccurately in place of *a* or vice versa). In the above table, we note that the most frequent type of inaccuracy uncovered in the written narratives was that of article omission. For instance, in contexts that required the use of *the*, beginners produced *the* with an accuracy of only 55.9%, a little more than a half of the NPs. As an alternative to *the*, they favored no article (37.5%), as in the following examples:

(19)   In that session, I had to prepare my graduate fashion show for 3 months. All my friends who prepared *_Ø_ graduate fashion show* with me and I thought, it's impossible. (B8)

(20)   Then, I got my new room, 327. During the week, I leaved my stuffs on the room, 301. I called it ground zero. When I got *_Ø_ new room*, Hotel started to clean _Ø_ *old room* up. (B9)

In (19), *graduation fashion show* was introduced in the first sentence and was subsequently mentioned in the following sentence, which requires the use of *the*. Also, in (20), *new room* was subsequently mentioned after the referent was introduced in the first sentence. Also, *old room* refers to Room 301, the room that the participant stayed in before the flooding occurred, which makes *old room* a definite NP. In both instances, the definite article was omitted.

Similarly, in contexts that required the production of *a*, beginners produced this form

accurately in only 38.6% of the cases. However, they opted to not use any article in 48.6% of the NPs instead of *a*. This pattern is indicative of the prevalent tendency to omit articles in written discourse as well as in spoken discourse. As many researchers suggest (e.g., Robertson, 2000; Thomas, 1989; Zegarac, 2004), the tendency to omit articles could possibly be due to an L1 transfer effect since Korean does not have a similar article system. The findings in the above table now confirm that article omission is also attested in a writing task even though sufficient time was provided. That is, beginners still tended to omit articles in their written narratives.

With regard to article over-generalization, *the* was employed inaccurately instead of *a* with a small but noticeable frequency of 7.9%. This suggests that beginners mistakenly judged indefinite contexts as definite contexts, and thus produced *the* instead of *a*, as in the following examples. In the following examples, *the* was used in contexts in which specific referents were introduced for the first time in the discourse, contexts which call for the use of *a*.

(21)   I participated in **the discussion competition* to get a prize. (B10)
(22)   We ate pasta in restaurant and then went to *the piano bar. (B2)

In this regard, several scholars (Ionin et al., 2004; Park, 2006; Thomas, 1989) found that L2 English learners tend to over-generalize *the* when noun referents were indefinite and specific. This finding thus provides evidence that beginner learners of English exhibit the same tendency to over-generalize *the* for specific, indefinite referents in the written production as well.

The next table examines the effect that the presence of adjectives in NPs had on the accurate use of the two articles.

Table 4. 18. The distribution of articles produced by beginners in the presence and absence of adjectives in NP

| Required article | Frequency of produced article | | | | | |
| | [+ Adjective in NP] ($N = 58$) | | | [- Adjective in NP] ($N = 223$) | | |
	the	*a*	omission	*the*	*a*	omission
the	53.8%	7.7%	30.8%	59.1%	0.7%	40.2%
a	4.4%	31.1%	62.2%	9.8%	43.9%	43.9%

$p < .0001$

In this table, we find that the beginners produced the definite and indefinite articles more accurately in contexts in which adjectives were not present in the NPs. For instance, *the* was produced with an accuracy of 53.8% in [+ Adjective in NP] contexts, but the accuracy for *the* increased six percentage points to 59.1% in [- Adjective in NP] contexts. A similar pattern was observed in the production of *a*. In NP contexts that contained adjectives, the accuracy in the written production of *a* was only 31.1%. On the other hand, in instances in which *a* was accompanied by an NP without an adjective, the accuracy increased to 43.9%.

In the case of *the*, the accuracy was higher in [- Adjective in NP] contexts but article omissions were also more frequent in [- Adjective in NP] contexts. These patterns seem to indicate that the presence of adjectives in NPs did not condition the use of the definite article in written production. Recall that in the oral narratives, the accuracy that was revealed for the use of *the* was higher in [- Adjective in NP] contexts and article omissions were more frequent in [+ Adjective in NP] contexts. As such, in the section on oral narratives, a working hypothesis was proposed, which posited that the lower accuracy obtained for *the* and *a* in [+ Adjective in NP] contexts in oral narratives was attributed to the beginners' tendency to omit articles in instances that a pre-noun position was occupied by a lexical item. In other words, when two competing forms can be inserted before a noun, beginners tend to omit an article and express an adjective, a

form that has more semantic significance than an article. The hypothesis was formulated on the assumption that L2 speakers of English have limited cognitive resources available for spontaneous language production and that they may choose to omit articles and express adjectives when both occur in NPs.

In writing, however, constraints on cognitive processing are assumed to be relatively small because participants were allowed to take as much time as they needed to complete the task. As such, in writing, participants could plan ahead the types of nominals they wish to use. Even when they forgot to use articles, it is possible in writing for them to go back and insert articles later. In contrast, oral production involves linear processing of language and participants cannot go back and make revisions once they utter nominal expressions without articles.

In Table 4.18, we found that adjectives did not seem to influence the written production of *the*, since almost half of the instances were accurate (53.8%) in NPs with adjectives. In turn, the effect of the presence of adjectives was more evident on the accurate written production of *a*. For instance, the indefinite article was omitted in 43.9% of the NPs without adjectives, but the rate of *a* omission increased to 62.2% in instances in which adjectives were present, as in the following examples.

(23)   Suddenly I faced *∅ big problem (B7)
(24)   I met a professor who was preparing to open *∅ new alternative school (B10)
(25)   Suddenly she could feel *∅ strong hit (B3)

In the above examples, the indefinite article *a* was omitted in instances in which adjectives modified nominals. The more frequent omission of *a* in [+ Adjective in NP] contexts was also observed in the oral discourse. This finding suggests that the indefinite article is not a stable part of the beginners' implicit knowledge of articles and thus is more easily affected by a

144

variable such as the presence of adjectives than the definite article. One possible explanation for the effect of adjectives in NPs on the accurate written production of *a* is the lack of semantic transparency that characterizes the indefinite article. These observations seem to suggest that the beginners remain unaware of this form's importance in writing and its appropriate use regardless of planning and time provided to check the written discourse.

In the next table, the effect that NP contexts had on the accurate use of articles is examined. As mentioned earlier in the section on article use in oral narratives, NP contexts were coded into four types. The definite article *the* can be used in anaphoric definite contexts (i.e., second and subsequent mentions of an already introduced referent, as in *I bought a book yesterday. The book was on sale.*), and associative definite contexts (i.e., the definiteness of a noun referent is determined by contextual factors such as common knowledge or physical presence of a referent, as in *I bought a book yesterday. The cover was red.*). The indefinite article *a* can be used in referential contexts (i.e., first mention of a referent that is specific yet unknown to the hearer, as in *I bought a gift for you.*) and non-referential contexts (i.e., a noun referent that does not refer to any specific entity, as in *I feel lonely. I need a friend.*). The following table shows the frequencies of articles produced in each NP context type.

Table 4. 19. The frequency of articles used in the written narratives according to NP context

Required article	*N*	NP context	Frequency of produced article		
			the	*a*	omission
the	142	Anaphoric definite	56.9%	3.0%	40.1%
		Associative definite	58.4%	0.0%	41.6%
a	132	Referential indefinite	10.7%	46.4%	42.8%
		Non-referential indefinite	6.6%	36.8%	56.6%

p < .0001

In this table, we find that the accurate use of the definite article *the* in written discourse

was slightly higher in associative definite contexts. For instance, when *the* was the required article in a context and the use of this form was related to an NP that was previously mentioned, beginners produced the form with an accuracy of 56.9%, a little over a half of the times they had to produce the article. In turn, when the production of *the* was related to an associative definite context, the accuracy was slightly higher (58.4%).

With regard to the indefinite article, the beginners produced *a* more accurately in referential indefinite contexts (41.3%) than non-referential indefinite contexts (36.8%), a slightly higher frequency. In turn, the beginners used *the* instead of *a* more frequently in referential indefinite contexts (10.7%) than in non-referential contexts (6.6%). Similar to the findings obtained with the oral narratives, the beginners may have perceived referential indefinite contexts as definite contexts due to the specificity of a noun referent and favored *the* in place of *a*. The following examples exemplify their inaccurate uses.

> (26) I saw **the advertisement* which announced that new film school would open (B10)
> (27) I participated in **the discussion competition* to get a prize. (B10)
> (28) We ate pasta in restaurant and then went to **the piano bar*. (B2)

In these examples, referents were specific to the writers, but the readers were not familiar with them because the referents were introduced for the first time in the written narratives. In English, a specific referent is considered indefinite when it is not known to the hearer or the reader. Therefore, in these instances, the indefinite article *a* is expected. However, there is a tendency for L2 English learners to confuse specific indefinite referents for definite referents due to their semantic value of specificity. For example, Thomas (1989) asked 30 adult participants to perform a paired story-telling task and found that *the* was over-generalized in contexts in which specific referents were mentioned for the first time. In the case of this study, the beginners also over-

146

generalized *the* in referential indefinite contexts, as evidenced in the examples above.

In summary, beginners employed English articles in the written narratives in the following patterns:

(1) *The* was used with higher accuracy than *a*.

(2) When *the* and *a* were not produced accurately, a tendency to omit them was evidenced.

(3) *The* and *a* were produced more accurately in [-Adjective in NP] contexts.

(4) *The* was produced with similar rates of accuracy in anaphoric definite contexts and associative definite contexts.

(5) *A* was produced more accurately in referential indefinite contexts.

In this section, we examined how the beginners used articles in their written narratives. The next section discusses the findings on article use in the written narratives produced by the participants with intermediate language proficiency.

4.2.3. Intermediates: patterns of article use in written narratives.

This section discusses how the intermediates used articles in written narratives. This part of the Chapter contributes to addressing the first research question with regard to patterns of article use in different proficiency groups by focusing on the intermediate proficiency level in the written discourse. Following the same organization of the previous segment on the beginners' use of articles in written narratives, Table 4.20 first presents the distribution in the use of determiners by the intermediate participants.

Table 4. 20. The accuracy with which intermediates produced determiners in written narratives

Required determiner	*N*	Accuracy (%)
Definite article (*the*)	144	72.2
Indefinite article (*a*)	106	70.8
Quantifiers (e.g. *three*, *many*)	49	98.0
Possessives (e.g. *my*, *his*)	104	95.2
Demonstratives (e.g. *this*, *that*)	20	90.0
Others (e.g. *each*, *another*)	7	85.7
Bare NP	135	76.3

In Table 4.20, we find that the accuracy with which intermediates used articles was consistently high for the two articles. For instance, the accuracy for *the* was 72.2%, while the accuracy for *a* was 70.8%. Thus, the intermediates employed articles with higher rates of accuracy compared to beginners (70.8% and 38.6%, respectively) in written discourse. This observation suggests that intermediates in this study have reached a higher threshold of accuracy, one in which they are able to better identify the appropriate contexts for which articles need to be employed in writing.

With respect to the other types of determiners, the table reveals a relatively high accuracy (i.e., quantifiers 98%, possessives 95.2%, demonstratives 90%, and others 85.7%). Thus, the intermediates in this table exhibit a similar tendency as beginners with regard to the use of other types of determiners. As suggested in the beginner group, this tendency may be conditioned by the presence of equivalent determiner forms in Korean.

With regard to the accurate production in bare NP contexts, the intermediate group exhibited a different tendency. For instance, they exhibited an accuracy of 76.3%, which is lower than the beginner group (86.4%). The lower accuracy in contexts for bare NPs at the intermediate level may be related to the phenomenon called *U-shaped behavior* (Kellerman 1985, as cited in Ellis, 1994), which is concerned with positive L1 transfer. Ellis (1994) explains that "learners

may sometimes pass through an early stage of development where they manifest correct use of a

target-language feature if this feature corresponds to an L1 feature and then, subsequently,

replace it with a developmental feature before finally returning to the correct target-language

feature" (p. 303). As shown in the beginner section, the high accuracy in bare NP contexts in the

beginner group can be possibly attributed to the facilitative effect of L1. At the intermediate level

of proficiency, participants may have been developing and testing hypotheses about English

articles, which resulted in lower accuracy in bare NP contexts.

The next table illustrates the articles that the intermediates produced in their written

narratives with respect to the required articles. This table also shows the articles employed by the

intermediates when they did not produce the required articles. The *others* category in the table

includes determiners other than articles (i.e., possessives, demonstratives, and quantifiers).

Table 4. 21. The intersection between required articles and produced determiners in intermediates' written narratives

Required article	Frequency of produced determiner				
	the	*a*	omission	others	total
the	72.2%	4.2%	22.2%	1.4%	100.0% (N = 144)
a	0.9%	70.8%	27.4%	0.9%	100.0% (N = 106)

p < .0001

Table 4.21 shows that the intermediates produced articles with relatively good accuracy.

For instance, in contexts that required *the*, the intermediates produced the required form

relatively accurately in 72.2% of the NPs. Also, in contexts that required *a*, the intermediates

used *a* with a similarly high rate of accuracy (70.8%). When they did not produce the required

articles, they showed a preference for no article (22.2% when *the* was required, and 27.4% when

a was required). In other words, most of the inaccurate article instances were article omissions in

writing. Researchers have documented that L2 learners of English whose L1s do not employ

articles continue to exhibit a tendency to omit an article even as they become more proficient in

English (Park, 2006; Robertson, 2000; Thomas, 1989). Even though the intermediates were able

to produce the articles very accurately in the written narratives, they exhibited a residual

tendency to omit articles, even though these data point to relatively low frequencies.

The next table examines whether the variable of the presence of adjectives in NPs has

any effect on the accurate use of articles.

Table 4. 22. The distribution of articles produced by intermediates in the presence and absence of adjectives in NP

Required article	Frequency of produced article					
	[+ Adjective in NP] ($N = 63$)			[- Adjective in NP] ($N = 187$)		
	the	*a*	omission	*the*	*a*	omission
the	72.0%	0.0%	28.0%	72.9%	5.1%	21.2%
a	2.7%	67.6%	29.7%	0.0%	72.5%	26.1%

$p < .0001$

In Table 4.22, we find that the presence of adjectives in NPs does not influence the

accurate written production of the definite article *the*. For instance, *the* was produced accurately

in 72.0% of [+ Adjective in NP] contexts and 72.9% in [- Adjective in NP] contexts. Meanwhile,

the intermediates' ability to produce the indefinite article *a* accurately was influenced more by

the presence of adjectives in NPs than the definite article. For instance, the accuracy for *a* was

67.6% in [+ Adjective in NP] contexts, but the accuracy increased to 72.5% in NP contexts that

did not contain adjectives. The considerable difference in frequencies in the accuracy for *a*

between the two contexts suggest that intermediates seem to omit the indefinite article more

readily when NPs contain adjectives, as in the following examples.

(29) It was *∅ *really hard time* for me that time (I4)
(30) At that time, I made *∅ *Korean girl friend* (I7)
(31) I had to take *∅ *long walk* to go to Harvard (I9)
(32) I will apply to *∅ *big company* (I5)

In the above examples, *a* was omitted in instances in which adjectives were used to modify indefinite noun referents. The omission of *a* suggests that the intermediates have not reached a level of stability in the use of the indefinite article. Indeed, intermediates produced *a* with an accuracy of only 34.3% in the oral narratives. The effect of adjectives on the accurate written production of *a* is a further indication of the on-going development of the knowledge of the indefinite article by the intermediates.

Finally, the effect of the variable of NP contexts on the accurate use of articles is discussed. This analysis shows whether the article use by the intermediates was conditioned by distinct contexts in which particular articles are required.

Table 4. 23. The frequency of articles used in the intermediates' written narratives according to NP contexts

Required article	N	NP context	Frequency of produced article		
			the	*a*	omission
the	140	Anaphoric definite	84.8%	0.0%	15.2%
		Associative definite	69.1%	5.3%	25.6%
a	104	Referential indefinite	3.3%	66.7%	30.0%
		Non-referential indefinite	0.0%	73.0%	27.0%

p < .0001

Table 4.23 reveals differences in the accuracy for *the* and *a* depending on NP contexts. For instance, when *the* was the required article in a context and the use of this form was related to the NP that was mentioned previously, intermediates had a tendency to use *the* with a high accuracy of 84.8%. This observation suggests that intermediates were able to recognize instances in which a previously introduced noun referent was mentioned subsequently and provide *the* accurately in their written narratives. Alternatively, when the use of *the* was related to an associative definite context, accuracy was slightly lower at 69.1%.

With regard to the use of *the* in these two NP contexts, Kim & Ahn (2003) showed that Korean high school students seem to acquire the use of *the* in anaphoric definite contexts first. The researchers used an instrument that consisted of 87 sentences containing 58 deleted obligatory uses of *the*. Following Hawkins (1978), these uses were divided into four categories: textual use (i.e., when a noun is subsequently mentioned), situational use (i.e., when the referent of a first-mention noun can be sensed in the immediate situational context), cultural use (i.e., when a noun is a unique and well-known referent in a speech community), and structural use (i.e., when *the* is used along with certain types of modifiers). The participants were instructed to read the sentences and insert the definite article *the* wherever they believed necessary. The study revealed that the textual use, an equivalent of anaphoric definite contexts in this current study, was found to be easiest for the Korean participants. In comparison, cultural and structural use, which are referred to as associative definite contexts in this study, were most difficult. Thus, findings in Table 4.23 lend support to Kim and Ahn's findings.

With respect to the indefinite article, *a* was employed more accurately in non-referential indefinite contexts (73.0%) than in referential indefinite contexts (66.7%) in written discourse, a result similar to what was observed in the intermediates' oral narratives. Similar results were also reported by Ekiert (2010). Ekiert's study found evidence that "non-referential indefinite contexts were more readily identified by the participants as requiring some kind of indefiniteness marking in English" (p. 167). That is, referential contexts may allure learners to think that the definite marking is needed because of the specificity of these referents. In comparison, non-referential nominals do not refer to specific entities. In this regard, Ekiert (2010) notes that L2 learners receive early exposure to indefinite articles in many non-referential contexts (e.g., *I'm a student*) in an instructional setting. According to this perspective, intermediates in this study may have

152

produced *a* more accurately in non-referential indefinite contexts because they may have been exposed to the use of *a* in similar instances through explicit English language instruction.

In summary, the intermediates' use of articles in written narratives can be characterized in the following manner.

(1) The intermediates exhibited a more stable use of the articles in their written narratives.

(2) When the definite and indefinite articles were not produced accurately, article omission was preferred.

(3) The presence of adjectives in NPs showed a negative effect on the accuracy for *a*.

(4) The intermediates were able to identify a referent that was previously introduced and employ *the* with a high accuracy.

(5) Accuracy for *a* was higher in instances in which the indefinite article was used in non-referential indefinite contexts.

In this section, the findings point to how the intermediates command a much more stable and accurate use of articles in their written narratives compared to the beginner group. In the next section, findings on patterns of article use in the written narratives produced by the advanced participants are reported.

4.2.4. Advanced group: patterns of article use in written narratives.

This section presents the analyses of the English articles used by the advanced participants in the corpus of written narratives. Similar to previous sections, this section responds to the research question regarding the patterns of article use that emerge according to language proficiency. After examining these findings, the next section will provide a comparison of the

three groups.

To understand how advanced writers used articles, we first examine the frequencies of accuracy with which articles were produced in comparison with the accuracy for other determiners.

Table 4. 24. The accuracy with which the advanced produced determiners in written narratives

Required determiner	*N*	Accuracy (%)
Definite article (*the*)	218	76.1
Indefinite article (*a*)	178	75.8
Quantifiers (e.g. *three*, *many*)	57	100.0
Demonstratives (e.g. *this*, *that*)	34	100.0
Possessives (e.g. *my*, *his*)	171	98.2
Others (e.g. *each*, *another*)	14	92.9
Bare NP	140	82.9

Table 4.24 shows that the advanced participants employed the two articles at similarly high rates of accuracy. For example, the definite article was produced with an accuracy of 76.1% and the indefinite article, 75.8%. This finding suggests that the advanced participants do not show a tendency to fluctuate in their use of articles in written discourse as was the case in their oral discourse (accuracy of 76.2% for *the* and 55.1% for *a* in oral narratives).

The accuracy for other types of determiners was higher than the accuracy for articles. Specifically, the quantifiers were used with an accuracy of 100%, demonstratives, 100%, possessives, 98.2%, and others, 92.9%. This finding was also revealed in the spoken and written data for the beginner and the intermediate groups. Thus, based on these findings exhibited among beginners, intermediates, and advanced participants in both oral and written discourses, we can determine that these types of determiners do not pose a challenge to L2 English speakers of Korean L1. Therefore, further explanations about these forms such as demonstratives, possessives and quantifiers will not be addressed.

The advanced participants produced no article in bare NP contexts with an accuracy rate of 82.9%, a relatively high frequency. This trend seems to support the U-shaped behavior discussed in the intermediate section, a phenomenon of L2 learners exhibiting high accuracy for a language feature at the beginning due to positive L1 effect, lower accuracy in the middle of L2 development and returning to high accuracy at the advanced level. For instance, the accuracy in bare NP contexts was 86.4% in the beginner group, 76.3% in the intermediate group, and 82.9% in the advanced group. This finding lends support to claim that the beginners are under the positive L1 transfer effect. Intermediates, in turn, exhibited fluctuations and lower accuracy in these contexts, an indication of constant development and revisions of their interlanguage towards target-like use. Finally, the high accuracy in bare NP contexts among the advanced participants can be attributed to more target-like conceptualizations of articles and proper recognition of when no article is permitted.

The next table shows distributions of articles produced compared to articles required in given contexts, which illustrates the target-like and non-target-like production of articles in writing.

Table 4. 25. The intersection between required articles and produced determiners in the advanced participants' written narratives

Required article	Frequency of produced determiner				
	the	*a*	omission	others	total
the	76.1%	1.8%	17.4%	4.7%	100.0% (N = 218)
a	5.1%	75.8%	19.1%	0.0%	100.0% (N = 178)

p < .0001

As discussed in the previous section, we note again in this table that the advanced participants produced *the* and *a* with a high frequency of accuracy in their written narratives. For example, *the* was employed in 76.1% of the instances that required *the*. Similarly, *a* was

employed accurately in 75.8% of the instances. Thus, the advanced participants exhibited a stable and target-like use of both articles in their writing.

In turn, when the required articles were not produced, they tended to be omitted. The frequencies of omission were 17.4% for the definite article and 19.1% for the indefinite article. As mentioned in the section that reported on the intermediate group, the tendency to omit required articles seems to persist even in advanced levels of written production, a finding similar to Park (2006).

The next table examines whether the variable of the presence of adjectives in NPs has an effect on article use in the advanced group's written narratives.

Table 4. 26. The distribution of articles produced by the advanced participants in the presence and absence of adjectives in NPs

Required article	Frequency of produced article					
	[+ Adjective in NP] (*N* = 107)			[- Adjective in NP] (*N* = 279)		
	the	*a*	omission	*the*	*a*	omission
the	79.0%	4.7%	16.3%	80.0%	1.3%	18.7%
a	9.4%	62.5%	28.1%	2.6%	83.3%	14.0%

p < .0001

We find in this table that the advanced participants produced *the* with similar rates of accuracy in both [+ Adjective in NP] contexts and [- Adjective in NP] contexts, a finding similar to the intermediate group. For example, the advanced participants produced *the* with an accuracy of 79.0% in [+ Adjective in NP] contexts and 80.0% in [- Adjective in NP] contexts. Similar to the intermediate group, the stable production of *the* in both contexts suggests that the definite article seems to have stabilized in the advanced participants' written discourse.

While the presence of pre-nominal adjectives did not seem to influence the written production of *the*, it conditioned the accurate production of *a* in written narratives. For instance,

the advanced participants produced *a* with an accuracy of 62.5% in [+ Adjective in NP], but the accuracy increased greatly to 83.3% in [- Adjective in NP] contexts. Further, they tended to omit the indefinite article when adjectives were present in NPs, as in the following examples.

(33)   It was *ø rainy day when we had a trip (A2)
(34)   I am not really *ø good swimmer (A5)
(35)   I had same classmates for four years, which was *ø unique experience in my life (A9)
(36)   (…) give you *ø brief background history of my gained knowledge (A8)
(37)   It was *ø very harsh teaching or learning environment (A4)

These examples illustrate how the advanced participants favored no article instead of *a* when adjectives were present. Considering that *a* tends to be acquired later than *the* and that multiple elements of countability, number and definiteness need to be considered in arriving at the correct choice of article (Master, 1997), we may suggest at this juncture that the indefinite article has not become as stabilized in writing even at the advanced level.

Next, the written production of articles by the advanced participants is examined with regard to the variable of NP contexts. In the same manner with the other proficiency groups, the NP contexts were classified as the following: anaphoric definite contexts and associative definite contexts for the definite article *the*, and referential indefinite contexts and non-referential indefinite contexts for the indefinite article *a*. The frequencies of articles produced by the advanced group are shown for each NP context in the following table.

Table 4. 27. The frequency of articles used in the written narratives according to NP contexts

Required article	N	NP context	Frequency of produced article		
			the	*a*	omission
the	207	Anaphoric definite	79.6%	0.0%	20.4%
		Associative definite	80.3%	2.5%	17.1%
a	178	Referential indefinite	1.7%	68.3%	30.0%
		Non-referential indefinite	6.8%	79.5%	13.7%

p < .0001

In this table, we find that the indefinite article *a* was affected more by NP contexts than the definite article *the*. For instance, when the use of *the* was required for a referent that has been previously mentioned, the advanced participants accurately produced the article in 79.6% of the NPs, a relatively high frequency. Further, when the definiteness of a noun referent was determined by various contextual factors, *the* was also employed with an accuracy of 80.3%. This finding suggests that the advanced participants exhibit a stable production of the definite article in the written narratives regardless of NP contexts.

Unlike the tendency that was manifested in the stable written production of *the* in two distinct NP contexts, the advanced participants employed *a* with different rates of accuracy in referential indefinite contexts and non-referential indefinite contexts. For example, the accuracy with which the indefinite article was used was 79.5%, a strikingly high frequency. However, the same does not hold in referential indefinite contexts where we find that accuracy decreased to 68.3%. As mentioned in the intermediate section, Ekiert (2010) also found that the high-intermediate and advanced participants in her study used non-referential *a* more accurately than referential *a*. Further, the explanations provided after Table 4.26 suggest that the advanced participants may still be in the process of stabilizing the indefinite article in writing.

In summary, the advanced group used English articles in the written narratives in the following patterns:

(1) The articles were produced with relatively high accuracy.

(2) When *the* and *a* were not used accurately, writers tended to omit these forms.

(3) *A* was omitted more frequently when adjectives were present in NPs.

(4) *The* was produced with high rates of accuracy in both anaphoric definite contexts and

associative definite contexts.

(5) *A* was employed more accurately in non-referential indefinite contexts than

referential indefinite contexts.

In the next section, the patterns with which the participants used English articles are compared among the three proficiency levels. The comparative analyses are conducted to address the first research question of identifying patterns of article use along the proficiency continuum.

4.2.5. Group comparisons of article use in written narratives.

In this section, a comparison of how the participants in the three proficiency groups used articles in the written narratives is presented. The following figure illustrates accuracy patterns for the two articles (*the*, *a*) according to participants' proficiency.

Figure 4. 4. The accuracy in article use in written narratives according to proficiency
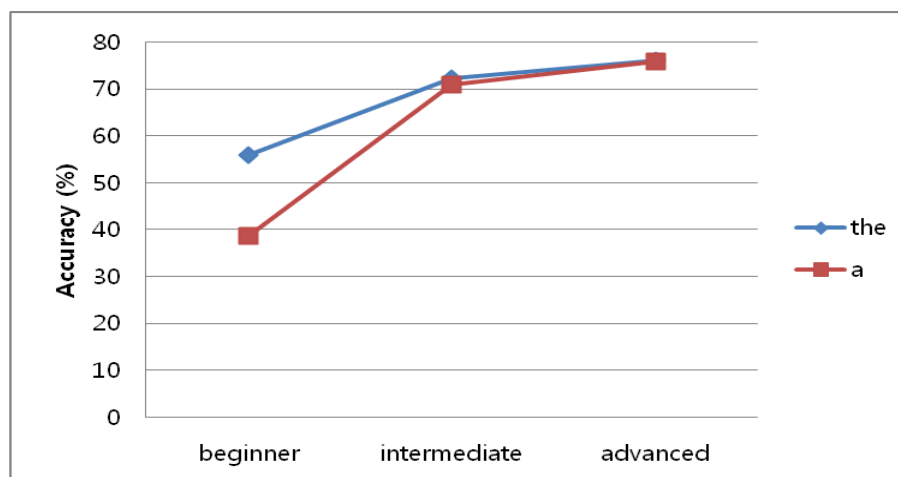
In this figure, several observations can be made with regard to the accuracy with which articles were used in writing by the three proficiency groups. First, we find that the beginners can be characterized with unstable uses of produced *the* and *a*. Specifically, they employed *the* with a higher accuracy than *a*. The higher accuracy for the definite article supports the documented

tendency for L2 learners of English to employ *the* more accurately than *a* at earlier stages of English language acquisition (Huebner, 1983; Master, 1987; Park, 2006; Parrish, 1987; Thomas, 1989). It seems that the beginner group was not able to accurately discern contexts in which *a* was required possibly due to their limited English language proficiency and the influence of L1. Master (1997) noted that "grammatical items that are comprised of multiple elements" (p. 220) are more difficult for language learners and that the elements of countability, number and definiteness need to be considered together to employ *a* accurately. Therefore, at the beginner level of proficiency, the use of *a* can be depicted as challenging even in written production. Conversely, the intermediates and the advanced participants show a gradual and stable increase in the use of *a* and *the* in written narratives.

A noteworthy observation can be made regarding this figure. There is a remarkable increase in the accuracy for articles from the beginner level to the intermediate level. Mizuno's (1999) cross-sectional study conducted four elicitation tests, including an essay writing test, on 2,000 Japanese learners of English. The study found that the improvement in article use in written essays was more observable between the beginner and the intermediate levels compared to improvement found between the intermediate and the advanced levels. Similar to Japanese, Korean does not utilize articles. Thus, the increases found in this study lend new support to Mizuno's findings with regard to the improvement as opposed to the stabilization of article use in writing.

Next, we compare patterns of article use in the written narratives produced by the three proficiency groups. Throughout this current study, instances of inaccurate article use are categorized into two: article omission and article over-generalization. Table 4.28 draws our attention to the frequencies of non-target-like article use in contexts in which *the* was required.

Table 4. 28. The frequencies of inaccurate article use in written contexts that required *the*

	Accuracy of *the*	Omission of *the*	Over-generalization of *a*
Beginner	55.9%	37.5%	1.3%
Intermediate	72.2%	22.2%	4.2%
Advanced	76.1%	17.4%	1.8%

In this table, we note that the most frequent type of inaccurate article use in writing is omission of *the* in all three groups (37.5% in the beginner group, 22.2% in the intermediate group, and 17.4% in the advanced group). The tendency for L2 English learners whose L1s do not have an article system, to omit an article has been consistently attributed to L1 transfer (Ekiert, 2010; Master, 1997; Park, 2006; Robertson, 2000, among others). Further, the findings in this table show that patterns of omission of *the* were evident not only in oral narratives but also in written narratives. Another observation from this table is that the frequencies decreased along the proficiency continuum. For example, the beginners omitted *the* in 37.5% of the contexts that required *the*, but the frequency decreased to 22.2% in the intermediate group and to 17.4% in the advanced group. As L1 influence is mostly related to article omission, this decrease suggests that L1 effect seems to diminish as proficiency increases.

With regard to instances of over-generalization, the participants used *a* inaccurately in instances in which *the* was expected with mixed frequencies. For instance, the beginners used *a* in only 1.3% of the contexts that required *the*. The frequency of over-generalizing *a* increased to 4.2% in the intermediate group but decreased to 1.8% in the advanced group, all relatively low frequencies. Ellis (1994) noted that the fluctuations in patterns of using certain linguistic features are often observed in the process of second language acquisition. In this regard, the fluctuations in the frequencies of *the* over-generalization across the proficiency levels may be conditioned by second language acquisition processes.

Similar to the previous table, the next table compares the frequencies of inaccurate

article use in instances in which the indefinite article was required.

Table 4. 29. The frequencies of non-target-like article use in written narrative contexts that required *a*

	Accuracy of *a*	Omission of *a*	Over-generalization of *the*
Beginner	38.6%	48.6%	7.9%
Intermediate	70.8%	27.4%	0.9%
Advanced	75.8%	19.1%	5.1%

In this table, we observe that the frequency with which the Korean learners of English

omitted *a* also decreased along proficiency levels. Specifically, the beginners omitted *a* more

frequently (48.6%), the intermediates, 27.4%, and the advanced participants, 19.1%. Similar to

the instances in which *the* was omitted in Table 4.28, this pattern is also indicative of a

decreasing L1 transfer effect along the proficiency continuum in writing. In other words, when

overall language proficiency increases, learners tend to produce *a* more accurately and the

frequency of omissions decreases.

Fluctuations are observed with regard to the over-generalization of *the*. For example, the

beginners used *the* inaccurately in 7.9% of the contexts that required *a*, but the intermediates,

only 0.9%. The frequency increased to 5.1% in the advanced group. Nevertheless, when

compared to patterns in oral narratives article production, the frequency with which the

participants over-generalized *the* was lower in the written production, a finding that was attested

in all proficiency groups. This means that with more time and attention available for review and

reflection on language in the writing task, the participants were able to avoid over-generalizing

the use of *the*.

Next, the effect of the presence of adjectives in the NPs on article use in the written

narratives is examined.

Figure 4. 5. The accuracy in article use in the presence and absence of adjectives in the NPs
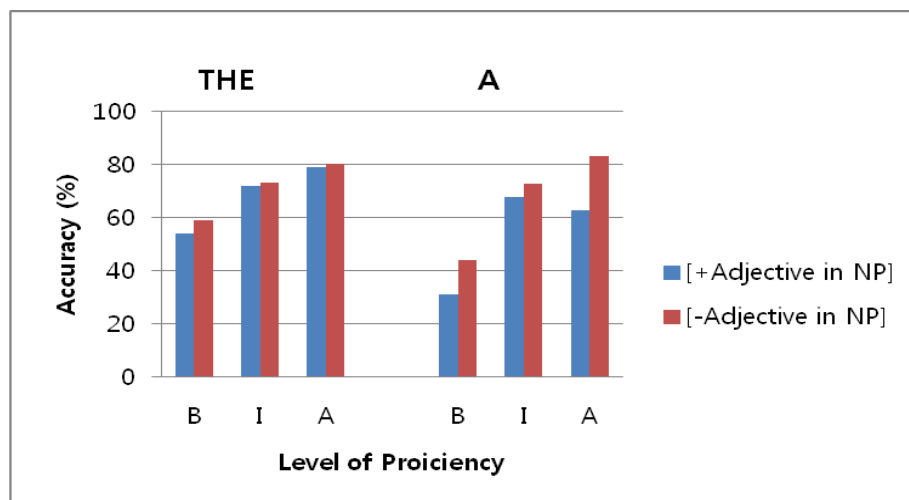
In Figure 4.5, at all levels of proficiency, the accurate production of *a* was affected more by the presence of adjectives in NPs compared to *the*. That is, the participants produced *the* within close ranges of frequencies although the beginners show a slightly lower trend. With regard to *a*, the figure points to higher levels of accuracy in [- Adjective in NP] contexts in all proficiency levels.

The above findings suggest that in writing, the presence of adjectives in nominals seems to have more effect on the accurate production of *a* than *the*. In comparison, Trenkic (2007) reported that both articles were affected by the presence of adjectives in NPs. A possible explanation for different findings between the two studies may be that Trenkic (2007) employed a more controlled translation task while the written corpus in this study consisted of unguided written narratives. We can suggest here that in the genre of written narratives, writers tend to use the definite article as a communicative and linguistic device to maintain referential coherence, which may in part explain why the definite article was less affected by adjectival modifications

than the indefinite article in written narratives. In turn, since the written production of *a* seems to be more challenging than the written production of *the*, the added cognitive burden of expressing adjectives in NPs may condition learners to omit *a* more frequently in [+Adjective in NP] contexts than *the*.

In the next section, I discuss how the variable of NP contexts affect accuracy. Figure 4.6 compares the accuracy for *the* in two NP contexts: anaphoric definite contexts and associative definite contexts.

Figure 4. 6. The accuracy for *the* according to NP contexts
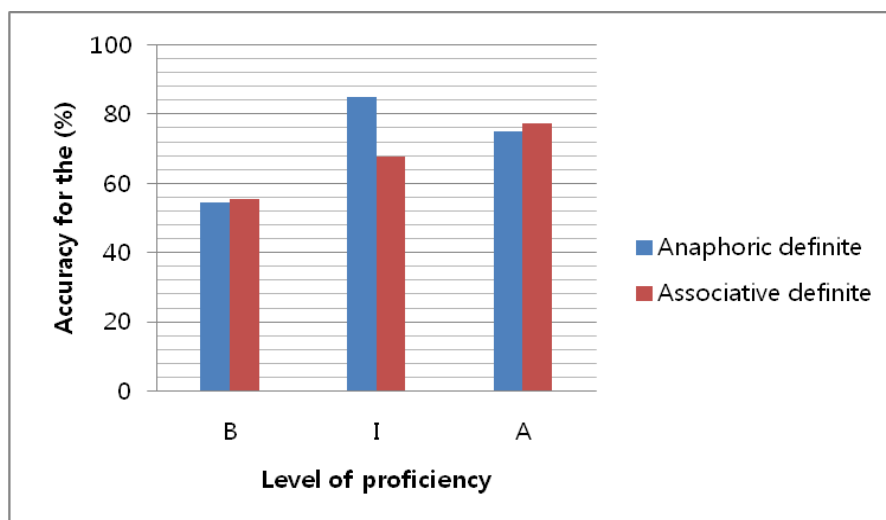
In this figure, we note that the accuracy for *the* in associative definite contexts increased gradually with proficiency, which is an expected finding as accuracy for most linguistic forms in L2 tends to increase as proficiency improves. However, the accuracy for the use of *the* in anaphoric definite contexts was found highest at the intermediate level. In a study on articles used by Korean learners of English in written composition, Park (2006) also found that the intermediate group produced the least article errors among the three proficiency groups. It is plausible to argue that the intermediate participants were sensitive to language rules and most

engaged in developing, testing and revising hypotheses on article use. Also, at the intermediate level, more monitoring and reflection of the language production may have occurred. Likewise, it seems that intermediate participants in this study were aware of the anaphoric use of the definite article and successfully tracked anaphoric chains of noun referents as they were engaged in producing written narratives. As a result, the intermediates were able to employ *the* very accurately in anaphoric definite contexts. In the case of the advanced group, *the* was produced with similar rates of accuracy in both NP contexts, a finding that indicates stability in article production in writing.

By comparison, the process of determining definiteness of nominals in associative definite contexts is far more complex than anaphoric definite contexts. In anaphoric definite contexts, *the* was used in instances in which the same NP that was previous introduced to the discourse is subsequently mentioned, as in *I bought a hat and a shirt today, but I will return the shirt tomorrow*. In comparison, associative definite contexts include a wider array of contexts in which nominals are considered definite. The following are several examples of the use of *the* in associative definite contexts.

(38) We went to a nearby beach. *The sand* was hot because it was early afternoon. (definite by association)
(39) What time does *the post office* open? (definite by assumed common knowledge)
(40) Can you bring me *the book*? (definite by the physical presence of the referent)
(41) The astronauts finally landed on *the moon*. (definite by the conventionally assumed unique referent)

Given the complexity in acquiring a variety of associative definite contexts that require not only structural but also cultural and pragmatic knowledge, L2 learners of English at lower proficiency levels may struggle with these contexts. However, the competency in the target-like use of *the* in associative definite contexts tends to improve along the proficiency continuum as

suggested by the above figure.

Following Figure 4.6, the next figure compares the accuracy for *a* in two NP contexts: referential indefinite contexts and non-referential indefinite contexts.

Figure 4. 7. The accuracy for *a* according to NP contexts
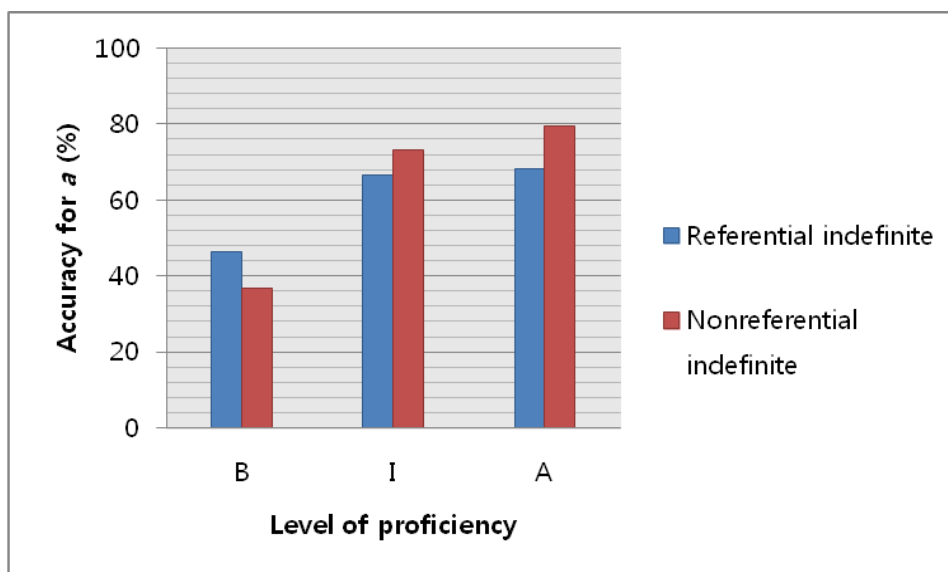
This figure shows that the beginner group employed *a* more accurately in referential indefinite contexts while the intermediate and the advanced groups used *a* more accurately in non-referential indefinite contexts. In general, the indefinite article *a* can be used in referential indefinite contexts (i.e., first mention of a referent that is specific yet unknown to the hearer as in *I bought a gift for you*), and non-referential indefinite contexts (i.e., a noun referent that does not refer to any specific entity as in *I feel lonely. I need a friend*). Ekiert (2010) found that the accuracy for non-referential *a* was higher than referential *a* in both the intermediate and advanced groups. Similar findings were obtained in this current study as the indefinite article was employed more accurately in non-referential contexts by the intermediate and advanced participants. However, the beginners exhibited a different pattern, an indication that they are still

at the early stage of second language acquisition.

The next section summarizes key findings of how participants used articles in the written narratives according to proficiency.

4.2.6. Summary of key findings: article use in written narratives.

In this section, patterns of article use in the written narratives were examined in three proficiency groups for the purpose of addressing the first research question that aimed to identify how articles are produced at different levels of proficiency. Specifically, articles produced by the participants were analyzed with regard to the variables of overall accuracy, patterns of non-target-like article use, the presence of adjectives in NPs, and the categories of NP contexts. Key findings can be summarized as the following:

(1) There was a positive correlation between accuracy for article use and proficiency levels in the written narratives.

(2) Rates of accuracy in the use of articles by the intermediate group approximated those of the advanced group in the written narratives.

(3) The most frequent kind of non-target-like article use was article omissions.

(4) The variable of the presence of adjectives in NP affected the accuracy in the use of *a* considerably more than the use of *the* in the written discourse.

(5) With regard to the variable of NP contexts for *the*, the accuracy in the use of *the* increased steadily along the proficiency continuum in associative definite contexts, while the accuracy in the use of *the* was highest in the intermediate group in anaphoric definite contexts.

(6) Between the two NP contexts for *a*, beginners produced *a* with higher accuracy in

referential indefinite contexts, while intermediates and advanced learners produced

a with higher accuracy in non-referential indefinite contexts.

The patterns of article that emerged in the analysis of the written corpus were summarized in the above. The extensive discussion of these findings in the context of the previous research is provided in Chapter VI. In the next section, patterns of article production are compared between the oral and the written discourses.

4.3. Comparisons of Oral and Written Production of Articles

4.3.1. Introduction.

In the previous sections, patterns of article use were examined in the oral narratives and the written narratives, respectively, and the focus of the analyses was placed on how distinct proficiency levels may affect the use of articles in spontaneous oral and written productions. In this section, we investigate the ways articles were employed in the oral narratives compared to the ways they were used in the written narratives. The purpose of the comparative analyses is to address the second research question that is concerned with comparing patterns of article use between the two modes of language production: spontaneous speech and planned writing.

Compared to the oral narrative task, the written narrative task may have offered advantages to the participants in terms of fluency, accuracy and complexity in the output. In particular, research has shown that accuracy is enhanced with on-line planning (i.e., within-task planning) as opposed to pretask planning (i.e., Ellis and Yuan, 2004; Meraji, 2011). For example, Ellis and Yuan (2004) studied the effects of three types of planning conditions (pretask planning, unpressured on-line planning, and no planning) on written narratives produced by Chinese learners of English. The results showed that unpressured on-line planning provided better

opportunities for monitoring, thus leading to increased accuracy in the written texts. Likewise, in this current study, the participants were instructed to spend as much time as they liked in producing written narratives. Therefore, we may hypothesize that the overall accuracy in the use of English articles would be higher in the written narratives compared to the oral narratives.

The positive effects of the affordance of time for monitoring and self-correction on accuracy improvement may, however, depend on the amount of explicit knowledge on English articles. In other words, if the knowledge on the English article system assumed by L2 writers is not complete and native-like, they may not be able to notice certain types of inaccurate article use and make appropriate revisions even though there is plenty of time available. On the other hand, the advanced participants might benefit more from the opportunity for on-line planning compared to the beginners because they have a better understanding of the article system. Furthermore, depending on the types and levels of knowledge of the article system, patterns of article use may vary.

In the following sections, I will examine whether the participants used articles more accurately in the written narratives than in the oral narratives as hypothesized. Results of the comparative analyses in the oral and written corpus are reported for individual proficiency groups. The final segment provides a summary of the findings.

4.3.2. Beginners: patterns of article use in oral and written narratives.

This section discusses how beginners used articles in oral narratives and written narratives. First, the accuracy in the use of articles by the beginner group is compared between the two discourses to understand the relationship between article accuracy and production mode.

Figure 4. 8. Accuracy in the use of articles by the beginners in oral narratives and written narratives
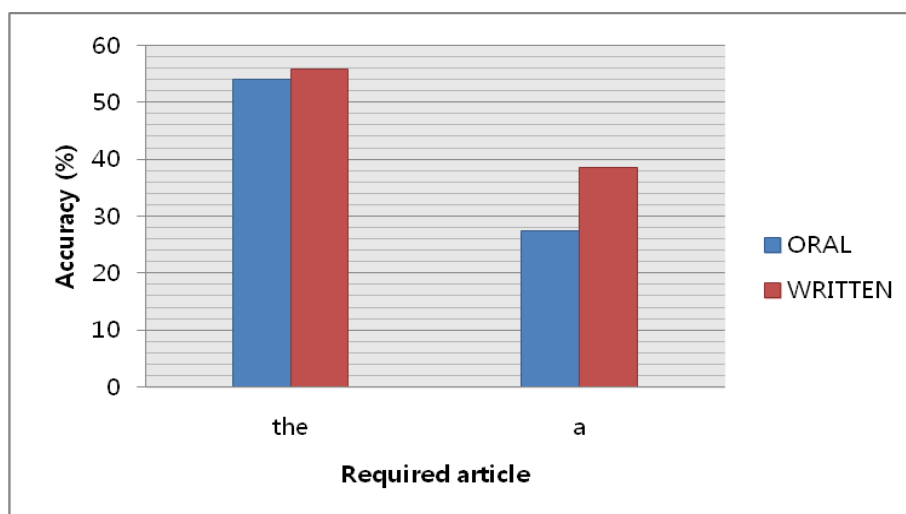
In Figure 4.8, we find that in general, the beginners employed both articles more accurately in writing than in speaking. This observation supports the hypothesis that articles would be used more accurately in the written text. Because writing is a planned discourse and speaking is not, participants seem to have been able to pay more attention to article use in their written narratives. Between the two articles, we can see that an oral versus written discourse shows distinct rates of accuracy. For example, the same beginners used *a* much more accurately in writing than in speaking, but they employed *the* with similar rates of accuracy in speaking and writing. This result suggests that the indefinite article is more challenging for the beginners to produce accurately in spontaneous production compared to the definite article because the accurate production of *a* involves correct identification of countability and noun numbers of nominal referents.

Indeed, concerns related to countability have been documented as sources of difficulties with regard to article production. For example, Yoon (1993) conducted a countability judgment test and a cloze test to 27 Japanese learners of English and 31 native English speakers. Along

with Korean, Chinese and many other languages that do not employ articles, Japanese does not explicitly mark noun numbers unless deemed necessary for communication. Yoon found that perception of countability of English nouns varies greatly by learners and that learners did not insert articles correctly when they perceived countable nouns as uncountable. This tendency revealed in Yoon's study suggests, therefore, that countability judgment is difficult for L2 learners of English from these L1 backgrounds.

The accurate perception of countability becomes more challenging in spontaneous speech when learners are cognitively burdened with organizing thoughts and structuring phrases and sentences to narrate a story. In writing, however, learners are afforded more time and attention to identify countability of nominals and employ appropriate articles. This difference may only partially account for the higher accuracy in the use of *a* in writing.

Since the overall rates of accuracy for the two articles were compared between speaking and writing, we now turn to particular variables that influence patterns of article use. In sections 4.1 and 4.2, we observed that the participants exhibited a tendency to omit articles more frequently when adjectives were present in NPs. Thus, "the asymmetry in article production in adjectivally modified and non-modified nominal contexts" (p. 309) reported by Trenkic (2007) was attested in both spoken and written production. Trenkic also stated that "the presence of an adjective increased the likelihood of the article drop in all learner groups, irrespective of the overall accuracy rates or the rate of omissions" (p. 308). The following table compares rates of article omission when NPs contained adjectives in oral narratives and written narratives produced by the beginner participants.

Table 4. 30. Frequency of article omissions by the beginner group in the oral and written narratives when adjectives were present in NPs

Discourse	Omission of *a*	Omission of *the*
Oral Narrative	73.2%	50.0%
Written Narrative	62.2%	30.8%

In this table, we note that the beginners omitted articles more frequently in oral narratives when nominals were modified by adjectives. For instance, they omitted *a* with a frequency of 73.2% in the oral narratives and with a lower frequency of 62.2% in the written narratives. Similarly, *the* was omitted in 50% of the [+Adjective in NP] contexts in the oral narratives, but the frequency decreased to 30.8% in the written narratives. These findings suggest that beginners' use of articles was less affected by the presence of adjectives in writing possibly because writing was not a spontaneous task but a planned activity. In speaking, the amount of an L2 speaker's cognitive and attentive resources directed to the employment of articles could be limited when adjectives needed to be expressed as well. In writing, L2 speakers may spend more time and attention to an accurate production of both articles and adjectives in NPs. This difference between oral and written discourses may possibly explain why the variable of the presence of adjectives in NPs had greater effect on article accuracy in speaking than in writing.

The beginner group also exhibited different patterns of article use in spoken and written discourses with regard to NP contexts in which *the* was required. Table 4.32 shows rates of accuracy in the spoken and written production of *the* in two NP contexts: anaphoric definite contexts (i.e., a previously introduced referent is mentioned subsequently) and associative definite contexts (i.*e.,* a referent takes the definite status due to its immediate situational context or its associations with a previously mentioned referent).

Table 4. 31. Accuracy in the use of *the* in two NP contexts across oral and written discourses

Discourse	Anaphoric definite contexts	Associative definite contexts
Oral Narrative	45.2%	63.0%
Written Narrative	56.9%	58.4%

We observe in this table that the beginners' production of *the* in anaphoric definite contexts was more accurate in the written narratives than the oral narratives (45.2% vs. 56.9%). In writing, all the utterances are available for review. Thus, it may have been less challenging for beginners to recognize anaphoric relationships and employ *the* for those nominals that had been previously introduced in the discourse.

In the associative definite contexts, the reverse pattern was evident. For example, in associative definite contexts, beginners employed *the* more accurately in oral narratives (63.0%) than in written narratives (58.4%). This finding appears to be contrary to the overall assumption that grammatical structures tend to be produced more accurately in written production than in oral production. Higher accuracy for *the* in associative definite contexts in oral narratives may be related to L2 learners' tendency to rely on formulaic chunks in speaking for the sake of fluency. Even though the current study did not code for the use of formulaic sequences[8], qualitative examination of the oral corpus produced by the beginners seems to suggest the reliance on formulaic language. Further, various instances of *the* in formulaic expressions fall mostly under the category of associative definite contexts in this study. The following examples are illustrative of such instances of *the* from beginners' oral narratives.

(42) In the morning, my cousin just drop over there. (B5)
(43) She picked up the phone and then she just said like that. (B5)

---

[8] Wray & Perkins (2000) defines the formulaic sequence as "a sequence, continuous or discontinuous, of words or other meaning elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar" (p. 1).

(44)   One day, I <u>went to the field</u> for training. (B6)
(45)   When I saw <u>the bible</u>, I afraid of God (B4)
(46)   I took the short film <u>for the first time</u> (B10)

As beginners utilized memorized formulaic expressions such as the above examples, they tended

to exhibit a more target-like use of the definite article, a possible explanation for the higher

accuracy for *the* in associative definite contexts in oral narratives compared to written narratives.

Next, we examine different patterns of article use in spoken and written production with

regard to NP contexts in which *a* was required. In Table 4.32, rates of accuracy in the oral and

written production of *a* are shown in two NP contexts: referential indefinite contexts and non-

referential indefinite contexts.

Table 4. 32. Accuracy in the use of *a* in two NP contexts across oral and written discourses

Discourse	Referential indefinite contexts	Non-referential indefinite contexts
Oral Narrative	29.0%	27.0%
Written Narrative	46.4%	36.8%

In this table, we note that beginners produced *a* more accurately in writing for both NP

contexts types for the indefinite article. In referential indefinite contexts, the accuracy for *a* was

29.0% in oral narratives but the accuracy was higher at 46.4% in written narratives. Similarly,

the indefinite article was used accurately in 27.0% of the spoken utterances and, in writing,

36.8%, a higher frequency. As discussed for Figure 4.8, writing seems to provide a more

favorable condition for L2 learners to identify contexts that require *a* and they employ the form

with better accuracy. Thus, here too we can observe a consistent pattern of higher rates of

accuracy in written discourse.

This segment reported distinct differences in patterns of article use between oral and

written narratives produced by the beginner group. The comparison of article use by

174

intermediates between oral and written narratives is reported next.

4.3.3. Intermediates: patterns of article use in oral and written narratives.

In this section articles produced by intermediates in oral narratives are compared with those produced in written narratives. As a first step, rates of accuracy in the use of articles are compared between the two modes of language production.

Figure 4. 9. Accuracy in the use of articles by the intermediates in oral narratives and written narratives
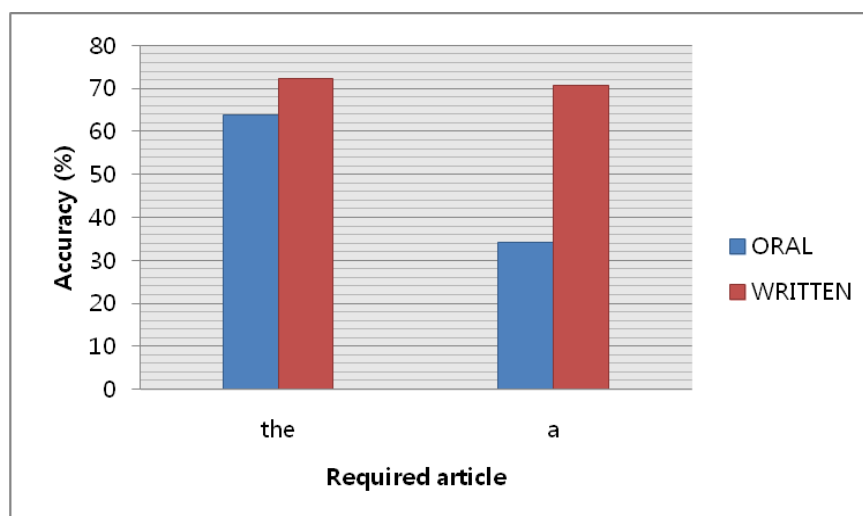
Similar to the beginner group, in general, the intermediates produced articles more accurately in writing than in speaking. The discourse effect was much more evident in the use of the indefinite article *a*. Specifically, the same participants employed *a* in written narratives with an accuracy nearly twice higher than that seen in oral narratives. This finding suggests that the intermediates have sufficient explicit knowledge to monitor their article use when they are given time and cognitive attention to review their writing. In case of beginners, the accuracy in the use of *a* was less than 40% even in written production, which is indicative of a lack of sufficient knowledge of articles to achieve higher accuracy. Comparatively speaking, intermediates were

able to employ *a* accurately in more than 70% of the instances in their written narratives. Thus, this finding point to conscious knowledge of English articles possessed by the intermediates as manifested in higher accuracy in writing. Also, the findings could suggest that the intermediates' knowledge of English articles is not completely implicit as their accuracy in the use of articles is still considerably lacking in oral production, in particular for the indefinite article.

The next variable to be examined is the effect of the presence of adjectives in NPs on article accuracy. In the beginner group, the tendency to omit articles in adjectivally modified NPs was observed more frequently in oral narratives. The following table shows whether this tendency was also evident in the intermediate group.

Table 4. 33. Frequency of article omission by the intermediate group in the oral and written narratives when adjectives were present in NPs

Discourse	Omission of *a*	Omission of *the*
Oral Narrative	76.7%	41.9%
Written Narrative	29.7%	28.0%

In this table, we find that rates of article omissions decreased considerably in the written narratives. Such discrepancy was more noticeable in the case of the indefinite article. Specifically, the intermediates omitted *a* in 76.7% of the instances in which adjectives were present in NPs in oral narratives, but they omitted the form in only 29.7% of such instances in the written discourse. The following examples are two utterances produced by the same intermediate participant (I3). ON refers to the oral narrative and WN, the written narrative.

(47)   (…) and we are *really good matched couple. (ON)
(48)   (…) and we were a good matched couple (WN)

In this pair of examples, these almost identical utterances were produced in almost identical contexts in the oral and written narratives. In speaking, the indefinite article was omitted while it

was produced accurately in writing. As discussed in the beginner section, the effect of the presence of adjectives in NPs seems to be greater in spontaneous speech possibly due to cognitive constraints experienced by L2 speakers. As the intermediates seem to have a conscious awareness of articles, they were able to better cope with the challenge of using adjectives to describe nominals and employing appropriate articles. This pattern may have also contributed to the higher accuracy for *a* in the intermediates' written narratives than in the oral narratives as shown in Figure 4.9.

We now turn our attention to the effect of different NP contexts on article use in two different discourses. Table 4.34 presents rates of accuracy for *the* in two definite NP contexts in the oral and written narratives.

Table 4. 34. Accuracy in the use of *the* in two NP contexts across oral and written discourses

Discourse	Anaphoric definite contexts	Associative definite contexts
Oral Narrative	58.5%	67.0%
Written Narrative	84.8%	69.1%

In the above table, we observe a dramatic increase in accuracy with which *the* was produced in anaphoric definite contexts across the discourses, while the accuracy in the use of *the* was similar between writing than speaking. For example, in oral narratives, *the* was employed accurately in 58.5% of the anaphoric definite contexts, but the accuracy increased to 84.8% in the written narratives, a remarkably high rate of accuracy. The following examples were produced by one intermediate participant (I6) in the oral and written narratives.

(49) let me call, like, one of doctor that I know, so we went to *<u>doctor's office</u> (ON)
(50) She called one doctor she's known and we went to <u>the doctor's office</u> (WN)

In example (49), the underlined phrase was produced without an article in the oral narrative and with the definite article in the written narrative. This is one of many instances in

which intermediates successfully kept track of anaphoric references and employed the definite article appropriately. This finding suggests that intermediates have an understanding of the role of the definite article in anaphoric expressions. The learners, however, might not have spontaneous access to this knowledge in spontaneous speech, thereby resulting in lower accuracy in the anaphoric use of *the* in oral narratives. Thus, the intermediates' explicit knowledge of articles seems to be more developed than the implicit knowledge. Intermediates can make use of their knowledge of articles in writing, a discourse that provides them with more time to make reflections and revisions, but the explicit knowledge may not be easily accessible in speaking when they have to produce in L2 in a spontaneous manner.

Next, we examine whether or not article accuracy was affected by the discourse of language production according to NP contexts for the indefinite article.

Table 4. 35. Accuracy in the use of *a* in two NP contexts across oral and written discourses

Discourse	Referential indefinite contexts	Non-referential indefinite contexts
Oral Narrative	31.3%	35.8%
Written Narrative	66.7%	73.0%

In both types of indefinite contexts, accuracy was higher in the written production. The accuracy in the use of *a* in referential indefinite contexts was 31.3% in speaking and much higher at 66.7% in writing. Similarly, *a* was produced accurately in 35.8% of the non-referential indefinite contexts in speaking and 73% in writing. These differences seem to reflect the overall tendency that target-like use of articles was more evident in the written narratives.

Much like the beginner group and as hypothesized, the written production of articles up to this point has tended to be more target-like than the production of articles in oral discourse. For instance, when nominals were modified by adjectives, articles were omitted more frequently

in speaking than in writing. In instances in which associative definite contexts were concerned, both articles were produced more accurately in all NP contexts when they were used in writing.

We now turn to the patterns of article use that the advanced group exhibited in speaking and writing in the next section.

4.3.4. Advanced group: patterns of article use in oral and written narratives.

This section addresses the effect of L2 production genres on article accuracy in the advanced group. It was assumed that articles would be used more accurately in writing because L2 learners can spare more time and attention towards grammatical accuracy when they are engaged in an untimed writing activity. First, rates of accuracy for article use are compared between the oral and written narratives in the following figure.

Figure 4. 10. Accuracy in the use of articles by the advanced participants in oral narratives and written narratives
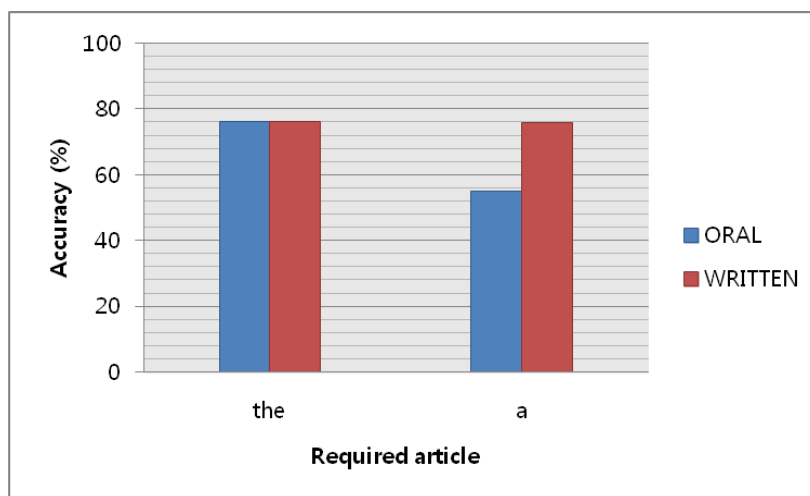
In this figure, we first note that the definite article was employed with similar rates of accuracy in writing and speaking. The observed stability in the use of *the* suggests that the participants in the advanced group have achieved a high level of proficiency with regard to the

definite article. Conversely, the use of *a* tended to be affected by the discursive modality in the

advanced group in much similar ways to those reported for the beginner and the intermediate

groups. Thus, we may infer that conceptual changes involving the indefinite article may be

delayed compared to the definite article for the advanced group.

With the overall comparison of article accuracy between speaking and writing, the next

table examines the rates of article omission according to the presence of adjectives in NPs.

Table 4. 36. Frequency of article omission by the advanced group in the oral and written
narratives when adjectives were present in NPs

Discourse	Omission of *a*	Omission of *the*
Oral Narrative	59.3%	36.4%
Written Narrative	28.1%	16.3%

The table shows that *a* was omitted in adjectivally modified NPs more frequently in oral

narratives (59.3%) than written narratives (28.1%), a tendency also observed in the beginner and

intermediate groups. The same pattern was observed with the definite article. Even though the

overall accuracy for *the* was stable across speaking and writing as shown in Figure 4.10, we still

find that accuracy in the use of *the* was mediated by the presence of adjectives in oral production

more so than in written production. When compared with written production, frequent omission

of articles in NPs containing adjectives in oral narratives does not seem to signal a lack of

linguistic competence regarding article use. This tendency is possibly due to performance

challenges that L2 learners experience when they attempt to produce longer phrases including

adjectives under time pressure in contexts of spontaneous production.

We now examine the effect of discourse with regard to article accuracy and the variable

of NP contexts.

Table 4. 37. Accuracy in the use of *the* in two NP contexts across oral and written discourses

Discourse	Anaphoric definite contexts	Associative definite contexts
Oral Narrative	84.1%	79.6%
Written Narrative	73.6%	80.3%

When *the* was required in anaphoric definite contexts for subsequent mentions, the form was produced more accurately in speaking than in writing. One possible explanation for this pattern is that an oral narrative task poses more communicative pressure on the narrator to maintain referential coherence. For instance, Tarone and Parrish (1988) compared article use in three different tasks: a written grammaticality judgment task, an oral interview, and an oral narration. The researchers also found that the definite article in anaphoric definite contexts was used most accurately in the oral narrative task. Similarly, the advanced participants in this study employed *the* more accurately in anaphoric definite contexts when the form was produced in oral narratives. In turn, rates of accuracy in the use of *the* in associative definite contexts were similar between the spoken and written discourses, an indication of stability in the use of *the* in these contexts.

Finally, the accuracy with which *a* was used is compared between the two genres of production according to two indefinite NP contexts.

Table 4. 38. Accuracy in the use of *a* in two NP contexts across oral and written discourses

Discourse	Referential indefinite contexts	Non-referential indefinite contexts
Oral Narrative	55.8%	55.6%
Written Narrative	68.3%	77.5%

This table shows that in both indefinite NP contexts, the indefinite article was used more accurately in written narratives than oral narratives. This tendency was observed in the other proficiency groups and may be related to the benefits of a written task, including more time to

review and review L2 production and more attention to grammatical accuracy by utilizing conscious knowledge of articles, among many other aspects of English.

4.3.5. Summary of key findings: article use in oral and written narratives.

Patterns of article use in oral and written discourses were compared at each proficiency level and the results were reported in the previous sections. The comparative analyses were aimed at responding to the second research question that focused on the differences and similarities in oral and written production of articles. In this section, a summary of key findings are provided. First, the beginner group exhibited the following characteristics with regard to article use in oral and written discourses:

(1) The beginner group produced *the* with similar rates of accuracy in both discourses.

(2) The beginner group produced *a* more accurately in written narratives than oral narratives.

(3) Beginners omitted articles more frequently in oral narratives than written narratives.

(4) In the oral narratives, beginners produced *the* more accurately in associative definite contexts than in anaphoric definite contexts, but the reversed pattern was observed in the written narratives.

(5) With regard to NP contexts for *a*, beginners employed *a* more accurately in referential indefinite contexts than non-referential contexts in both discourses.

Next, patterns of article use in the intermediate group were compared between the two discourses, and the following findings were obtained from the analyses.

(1) Intermediates produced both article forms more accurately in writing than in speaking, but the discourse effect was much more evident in the use of *a*.

(2) The intermediate group omitted articles more frequently in oral narratives than written narratives.

(3) With regard to NP contexts for *the*, intermediates produced the form more accurately in associative definite contexts than anaphoric definite contexts in the oral narratives. In the written discourses, however, the accuracy for *the* was considerably higher in anaphoric definite contexts than associative definite contexts.

(4) With regard to NP contexts for *a*, the intermediate group employed the form more accurately in referential indefinite contexts than non-referential contexts in both oral and written discourses.

Finally, a comparison was made between how the advanced participants used articles in oral narratives and how they did so in written narratives. The key findings can be summarized as the following:

(1) The advanced group employed *the* with similarly high rates of accuracy in speaking and writing.

(2) The indefinite article was produced more accurately in the written discourse.

(3) Article omissions were more frequent in the oral narratives than written narratives.

(4) The advanced group exhibited similar rates of accuracy in both NP contexts for *the* in both oral and written discourses.

(5) In both NP contexts for *a*, advanced participants employed the form more accurately in the written narratives than oral narratives.

In this chapter, patterns of article use were examined according to proficiency groups and modes of L2 production for the purpose of addressing the first two research questions concerning

patterns of article use according to proficiency levels and differences in article use between speaking and writing. In quantitative analyses, accuracy in the use of articles was cross-tabulated with variables such as the presence of adjectives in nominals and different NP contexts. Distinctive and progressive patterns of article use were observed at different levels of proficiency. Also, it was evident that article use was more accurate in writing possibly because writing is a planned discourse.

Chapter V. Results on Metalinguistic Knowledge of English Articles

The current study was designed to investigate both implicit and explicit knowledge of the English article system by learners of English from Korea whose L1 does not have a similar grammatical system. Implicit knowledge of English articles was explored by analyzing the participants' article use in spontaneous oral and written production of personal narratives according to proficiency levels. The patterns of article use in oral and written narratives produced by the three proficiency groups were reported in details in Chapter IV.

While implicit knowledge of a particular grammatical feature is manifested in its spontaneous use, explicit knowledge is typically measured by using carefully constructed instruments. In this study, two tasks were conducted to gather data on explicit knowledge of English articles: the grammaticality judgment task and the think-aloud task. Details on how these two tasks were administered are provided in Chapter III. The data obtained from the two tasks were analyzed qualitatively and quantitatively to evaluate the kinds of explicit knowledge of English articles displayed by the three proficiency groups.

This chapter reports results of data analyses from the grammaticality judgment task and the think-aloud task. The organization of this chapter is as follows. First, findings of the

grammaticality judgment task are presented. This is followed by the report and discussion of findings of the think-aloud task. More specifically, quantitative results of the fill-in-the-article task, which was implemented as part of the think-aloud task, are reported. Then, quantitative and qualitative analyses of the think-aloud protocol are reported to investigate reasons for article choices explicitly supplied by the participants at different proficiency levels to address Research Questions 3 and 4. Finally, a comparison is made between how the articles were used in the oral narratives of the selected participants and how they explicitly explained for their article choices in the think-aloud task. This comparison is made to address the fifth research question that intended to understand how metalinguistic hypotheses about articles may be operationalized in speech.

5.1. Performance on the Grammaticality Judgment Task

5.1.1. Results.

In this section, the results of the grammaticality judgment task by the thirty ELL participants and the five NES participants are presented and discussed. First, mean scores of the four groups are reported, followed by the results of independent samples *t*-tests to corroborate whether the mean scores of the groups were statistically significant. Next, scores obtained from the grammaticality judgment task are considered according to article types and relevant NP contexts to examine the aspects of article use that the participants exhibited knowledge.

As discussed in Chapter III, the grammaticality judgment task has been commonly used in SLA research as a test to evaluate knowledge of a particular grammar that second language learners possess. Mandell (1999) argued that data produced by grammaticality judgment tasks are reliable measures of linguistic knowledge. In this current study, a small-scale grammaticality

judgment task was administered as a supplementary source of data that could demonstrate the level of explicit knowledge of English articles by the participants. For this study, the participants were asked to judge if the use of underlined articles in 16 instances was grammatical. There were five instances of the definite article, five instances of the indefinite article, and six instances in which no article was required.

With regard to the overall performance of the grammaticality judgment task by the three proficiency groups as well as the NES group, the following table shows the mean accuracy scores and standard deviations of the four groups.

Table 5. 1. Group means and standard deviations of the grammaticality judgment task by proficiency groups

Proficiency group	Mean	Standard Deviation
Beginner group	74.35% (11.9/16)	7.48
Intermediate group	81.84% (13.1/16)	14.27
Advance group	77.47% (12.4/16)	11.10
NES group	96.24% (15.4/16)	5.59

Table 5.1 shows that the beginners' mean score was 74.35, the lowest average scores among the four groups. As expected, the NES group scored the highest. Among the learner groups, it is interesting to note that the mean score of the intermediate group is higher than the advanced group. This finding may possibly suggest that the intermediate participants have the highest level of explicit knowledge of English articles among the three proficiency groups. Further details about their metalinguistic knowledge will be presented in later sections of this chapter to capture their understanding of articles.

In order to find out if there are statistically significant differences among the proficiency groups with regard to the grammaticality judgment scores, a series of independent samples t-tests were carried out. The results are summarized in Table 5.2.

Table 5. 2. Independent samples *t*-tests for group comparisons of the grammaticality judgment scores

Group comparison	t	df	Sig. (two-tailed)
Beginner vs. Intermediate	-1.472	18	.158
Intermediate vs. Advanced	.765	18	.454
Beginner vs. Advanced	- .738	18	.470

The results of the independent samples *t*-tests show that there were no significant differences between proficiency groups with regard to the grammaticality judgment scores. That is, the three groups' performances on the grammaticality judgment task were similar. This result was unexpected, considering the fact that significant differences in the rates of accuracy were found among the three groups in their oral and written narratives. A possible explanation for undifferentiated mean scores among the three learner groups could be a consequence of the task itself. That is, the task may have been constructed in a way that did not require a more complex understanding of knowledge on articles. Namely, what was required to perform well in the grammaticality judgment task was basic rules about articles (e.g., the use of *the* for a subsequent mention and *a* for a new singular referent). Another plausible reason for the high scores achieved by all three groups is that participants at lower proficiency levels may indeed possess sufficient explicit knowledge of English articles to be able to perform well on this type of task but they may not have developed sufficient implicit knowledge to employ articles accurately in spontaneous language production. In this regard, the participants' metalinguistic knowledge of articles will be examined in more detail in the next sections when think-aloud data analyses are reported to enable the researcher to elucidate on their performance in the grammaticality task.

Next, scores on the grammaticality judgment task are divided for each article form and proficiency level in the following table. The 16 test items included five instances for the definite

article, five instances for the indefinite article, and six instances for bare NPs.

Table 5. 3. Accuracy in grammaticality judgments with regard to target articles

Target Form	Beginner	Intermediate	Advanced	NES
the	72.9%	74.3%	78.6%	100.0%
a	86.0%	100.0%	90.0%	96.0%
no article	62.5%	82.5%	60.0%	93.4%

The table shows that all three proficiency groups performed better with regard to the indefinite article *a*. For instance, the beginner group scored 86.0% in instances in which *a* was required, while the scores for *the* and bare NPs were 72.9% and 62.5%, respectively. Similarly, the intermediate group scored 100% on the questions about the indefinite article, but scored less for the definite article (74.3%) and bare NPs (82.5%). The advanced participants scored 90% accuracy for *a*, 78.6% for *the*, and 60% for bare NPs. The tendency for learners at all proficiency levels to perform better on items for *a* than with *the* is an unexpected finding because the target-like production of *a* was far less frequent than *the* in both oral and written narratives. The NES group achieved consistently high levels of accuracy for all target article forms.

The unexpected findings shown in the above table combined with the lack of statistically significant differences among the three proficiency groups (see Table 5.2) seem to suggest that the grammaticality judgment task employed in the current study is not an appropriate tool to determine levels of article knowledge according to proficiency. In this regard, Johnson et al. (1996 as cited in Mandell, 1999) reported that grammaticality judgment test responses from the adult L2 learners revealed a degree of inconsistency and concluded that the grammars of the adult learners are not fully determinate and that adult learners tend to rely on other strategies for their performance on grammaticality judgment tests. Similarly, Ellis (1991) documented within-subject inconsistencies in the judgments of the same sentences. The unexpected results of the

grammaticality judgment task employed in the current study may be partly attributed to these concerns of using grammaticality judgment tests with L2 learners.

Finally, the performance on the grammaticality judgment task is examined according to four NP contexts: anaphoric definite contexts and associative definite contexts for the definite article, and referential indefinite contexts and non-referential indefinite contexts for the indefinite article.

Table 5. 4. Accuracy on the grammaticality judgment task for different NP contexts

Target Article	NP context	Beginner	Intermediate	Advanced
the	Anaphoric definite	100.0%	100.0%	100.0%
	Associative definite	68.3%	70.0%	75.0%
a	Referential indefinite	82.5%	100.0%	95.0%
	Non-referential indefinite	100.%	100.0%	70.0%

With regard to the definite article, we note that the three groups scored 100% in anaphoric definite contexts. Rules associated with the use of *the* for anaphoric expressions are among the first rules that L2 learners of English receive from teachers and grammar books. Thus, in this grammaticality judgment task, all the participants successfully identified anaphora and judged the grammaticality of the definite article in these instances correctly.

In turn, from this table, we can observe mixed results with regard to the two NP contexts for the use of the indefinite article. The beginner group outperformed its counterparts in non-referential indefinite contexts (100%) than in referential indefinite contexts (82.5%). Conversely, the advanced group performed better in referential indefinite contexts (95%) than in non-referential indefinite contexts (70%). The intermediate group scored 100% in both NP contexts, which is suggestive of a high level of explicit knowledge of articles. Thus, we find in several instances that native-like behaviors (100%) are being exhibited among the groups with respect to

the contexts in which the article forms are used.

In this section, quantitative results of the grammaticality judgment task were reported. First of all, the NES group exhibited nearly perfect judgments of grammaticality in 16 instances of article use. The three learner groups were found to have performed with similar levels of correctness because the mean scores of the three groups did not differ statistically. Yet, the intermediate participants scored the highest among the learners.

Between the two articles, grammaticality judgment of various uses of *a* was more accurate than *the*, a tendency evidenced in all three groups. With regard to NP contexts, participants at all proficiency levels performed better in anaphoric definite contexts than associative definite contexts. The performance in two indefinite contexts fluctuated among the groups without clear patterns along the proficiency continuum. These findings are discussed in detail in the next section.

5.1.2. *Summary of key findings: grammaticality judgment task.*

In this current study, the grammaticality judgment task was administered as a supplementary measure to assess explicit knowledge of article use. The participants were asked to judge grammaticality of the articles used in 16 instances. As summarized in the above section, mixed results were obtained from the grammaticality judgment task, which contradicted the previous findings. Among the three learner groups, the intermediates outperformed the beginners and the advanced participants. However, the independent samples *t*-tests showed that there was an absence of statistical significance when the mean scores of the three proficiency groups were compared. Further, the variable of NP contexts for the indefinite article does not seem to have influenced the participants' performance.

190

One possibility to consider in interpreting the mixed results of the grammaticality judgment task is that participants might have used a variety of strategies in completing the task, strategies that this study cannot account for since they fall outside the scope of this study. For instance, Ellis (1991) reported that 21 adult advanced Chinese learners of English employed a variety of strategies in making judgments, including feeling, rehearsing, accessing explicit knowledge, using analogy, evaluating a sentence, and guessing. The scholar reported that (1) the learners in his study demonstrated considerable inconsistency in their judgments, (2) learners appeared to rely extensively on "feel" in order to make a judgment, and (3) learners were relatively unsuccessful in using explicit grammatical knowledge to make judgments. The scholar cautiously concluded that grammaticality judgment tasks need to be validated carefully before they can be used as a basis for making claims about competence. In this sense, the results of the grammaticality judgment task employed in this study can be considered inconclusive because the high scores produced by the beginners may not be a valid indication of their explicit knowledge of articles but other factors may have influenced the outcome.

In short, I consider the results of the grammaticality judgment task indeterminate because it is difficult to draw generalizations based on the mixed results of the task. The results can also be a result of the design of the grammaticality judgment task. Limitations with regard to this instrument are considered in detail in Chapter VII. From this point on, therefore, the results of the grammaticality judgment task will not be discussed further. Instead, the think-aloud data can now be considered of major importance in this current study because it allows us to gain a deeper understanding of their metalinguistic understanding of the article system. In the next section, results of the think-aloud task are reported.

5.2. Results of the Think-Aloud Task

The think-aloud protocol was designed to gain access to explicit and metalinguistic knowledge of English articles that the L2 speakers of English. For the purpose of making comparisons between the Korean participants and native English speakers (NES), the NES group also participated in the think-aloud task. This task was designed to address the research questions concerning (1) the kinds of metalinguistic awareness that Korean L2 speakers of English have at different English language proficiency levels and (2) how their explicit hypotheses about English articles are different from those of native English speakers. The metalinguistic task was also designed to enable the researcher to triangulate data obtained from the previous analyses conducted.

The think-aloud task was comprised of two components: the fill-in-the-article test and the think-aloud protocol. Participants were asked to take the fill-in-the-article test by supplying articles to 72 blanks provided in a written text. While they were making article choices for the blanks, they were also told to think aloud and make explicit their thinking process behind the article choices. These explanations were digitally recorded and coded. Naturally, the fill-in-the-article test produced quantitative data and the think-aloud protocol produced qualitative data.

Taking into account the kinds of data obtained from the think-aloud task and the various analyses conducted, this section is organized as follows. First, the results of the quantitative analyses of the fill-in-the-article test are reported and discussed. The segment on the fill-in-the-article test is followed by both quantitative and qualitative analyses of the data from the think-aloud protocol. The reasons that the participants orally expressed for their article choices were categorized and compared across proficiency levels as well as the NES group. This section concludes with a summary of key findings.

5.2.1. Performance on the fill-in-the-article test.

This section presents the results of a series of quantitative analyses performed on the results of the participants' fill-in-the-article test. First, the overall rates of accuracy obtained from the fill-in-the-article test of the four proficiency groups are presented and compared. Next, the accuracy rates are analyzed for the distinct article forms and then they are also compared across proficiency levels. Finally, accuracy rates in the fill-in-the-article test for the four NP contexts are compared across the proficiency groups.

The mean accuracy scores that the participants obtained on the fill-in-the-article test were calculated to understand how they performed on the test at each proficiency level.

Table 5. 5. Mean accuracy scores in the fill-in-the-article test by proficiency group

Proficiency Group	Mean (%)	Standard Deviations
Beginner Group ($N = 10$)	68.06	13.49
Intermediate Group ($N = 10$)	78.88	5.94
Advanced Group ($N = 10$)	80.12	7.27
NES Group ($N = 5$)	94.44	3.23

The table shows that the mean scores of the fill-in-the-article test increased along with the proficiency continuum. The mean accuracy score of the beginner group is 68.06%, while that of the intermediate group was higher at 78.88%. The advanced group scored 80.12% accuracy, which is slightly higher than that of the intermediate group. The mean accuracy of the NES group was the highest with 94.44% accuracy. We can observe that there is a considerable gap in the mean accuracy scores between the NES group and the advanced group. Butler (2002) also reported that there was a substantial gap in the mean scores of the native English speakers and those of the Japanese L2 learners of English in the fill-in-the-article test. In the current study, the

highest score among the advanced participants was still considerably lower than the lowest score among the NES participants. This observation seems to suggest that it is extremely difficult to achieve a native-like command of English articles for L2 learners of English, especially when their L1s do not employ articles, but the data shown in the table is also suggestive of this: As proficiency increased, improved uses of articles increased.

Next, independent samples *t*-tests were conducted to compare mean scores of two proficiency groups for statistically significant differences. The results are summarized in the following table.

Table 5. 6. Independent samples *t*-tests of mean scores on the fill-in-the-article test

Group comparison	t	df	Sig. (two-tailed)
Beginner vs. Intermediate	-2.321	18	.038
Intermediate vs. Advanced	-.417	18	.681
Beginner vs. Advanced	- 2.488	18	.023

p < .05

In this table, we can see that there is a statistically significant difference between the beginner group and the intermediate group with regard to the accuracy scores in the fill-in-the-article test. However, a significant difference was not found between the intermediate group and the advanced group. This finding suggests that the intermediate participants and advanced participants performed similarly on the fill-in-the-article test.

In the fill-in-the-article test, three choices were available for the participants to make: (1) *the*, (2) *a*, (3) $\emptyset$ [9]. For each target form, accuracy scores were calculated according to proficiency group. For reference, accuracy scores for individual participants are available in Appendix L.

---

[9] This symbol ($\emptyset$) was used in the think-aloud task to indicate an instance in which no overt article was required. The entire protocol is provided in Appendix J.

Table 5. 7. Accuracy in article selections in the fill-in-the-article test according to target articles and proficiency levels

Target article	Beginner	Intermediate	Advanced	NES
the	70.0	72.7	80.5	94.6
a	71.5	92.0	85.5	95.0
Ø	57.3	76.7	72.0	93.3

$p < .0001$

In this table, we note that the accuracy with which *a* was appropriately selected was higher than the other two choices (*the* and Ø) in all four proficiency groups. The rates of accuracy in instances for *a* were 71.5% in the beginner group, 92.0% in the intermediate group, 85.5% in the advanced group, and 95.0% in the NES group. This pattern is similar to what was observed in the grammaticality judgment task, for which accuracy for *a* was higher than accuracy for the in all proficiency groups. This finding suggests that participants may have the explicit knowledge required to accurately determine instances in which the indefinite article is required, but their implicit knowledge may not approximate the level of explicit knowledge. The implicit knowledge is more involved in the accurate production of the form in natural language use.

Another observation can be derived from this table. The accuracy in the selection of *the* increased along with proficiency while the accuracy in the selection of *a* and Ø fluctuated among the three participant groups. For instance, with regard to the definite article, accuracy increased steadily from the beginner level to the advanced level. Conversely, in instances in which *a* was needed, the intermediate group performed better (92.0%) than the advanced group (85.5%). Similarly, the intermediate group exhibited a slightly higher accuracy in the selection of Ø (76.7%) than the advanced group (72.0%). The high scores obtained on the fill-in-the-article test

suggest that the intermediate group may possess the highest level of explicit knowledge.

In the next table, the accuracy of article choices in the fill-in-the-article test was examined with regard to the variable of NP contexts in the four proficiency groups. In the current study, two NP contexts were identified for the definite article: 'anaphoric definite context' (i.e., when a referent is mentioned again after it has been introduced in the discourse), and 'associative definite context' (i.e., when the definiteness of a referent is determined by a number of contextual factors). Likewise, two NP contexts were identified for the indefinite article: referential indefinite context (i.e., the first-mention of a referent that is specific yet unknown to the hearer) and non-referential indefinite context (i.e., when a noun referent does not refer to any specific entity).

Table 5. 8. Accuracy in article choices according to NP contexts and proficiency levels

	NP contexts	Beginner	Intermediate	Advanced	NES
the	Anaphoric definite	76.7	93.9	86.1	98.9
	Associative definite	64.7	52.6	75.3	90.5
a	Referential indefinite	74.0	95.3	85.3	94.7
	Non-referential indefinite	64.0	82.0	86.0	96.0

p < .0001

Several observations can be made from this table. First of all, we note that between the two NP contexts for the definite article, all the groups scored with high rates in contexts in which the definite article was produced in anaphoric contexts (76.7% vs. 64.7% in the beginner group, 93.9% vs. 52.6% in the intermediate group, 86.1% vs. 75.3% in the advanced group, and 98.9% vs. 90.5% in the advanced group). As White (2010) points out, the employment of *the* in anaphoric definite contexts is the simplest article rule introduced in various ESL textbooks. Thus, learners can easily insert *the* in this explicitly article-oriented task when previously introduced referents are mentioned subsequently. The simplicity and clarity of the article rule applied for

anaphoric definite contexts may account for this stable outcome across the groups.

Another observation that can be made from this table is that the intermediate group exhibited a very high score in the use of the definite article in anaphoric contexts. In other words, the intermediates are approximating the performance of the NES group, 92.9% and 98.9%, respectively. Yet, in the category of associative definite contexts, the intermediates scored below all the groups (52.6%). In these associative definite contexts, even though nominal referents were not previously mentioned, they may take the definite article by association with previous text or due to contextual factors. For instance, one may say, *I opened <u>the fridge</u> but could not find anything to eat*, the definite article can be used for *fridge* in its first-mention when it is obvious to both the speaker and the hearer which fridge NP refers to. If a learner chooses *a* on the ground for its first-mention, it results in an inaccurate article selection. Reasons for the intermediates' non-target-like article choices in these contexts will be examined in detail in the next section.

With regard to the contexts for the indefinite article, mixed results emerged. In the beginner and intermediate groups, referential indefinite contexts were found to influence accuracy rather than non-referential indefinite contexts (74.0% vs. 64.0% in the beginner group, and 95.3% vs. 82.0% in the intermediate group). In other words, we find that participants at lower proficiency levels seem to be able to judge target-like use of *a* more accurately when the form was used for a specific referent than when it was for non-referential use. However, the advanced group and the NES group exhibited similar performances in the two indefinite contexts, a finding that points to the stability in their understanding of the use of *a* for the advanced group and the NES group.

In this section, results of the fill-in-the-article test were reported. It is noteworthy to mention at this juncture that the NES group did not score 100% on the test even though they

were recruited as native English speakers. Several possible reasons can be suggested for this observation. First, we need to acknowledge individual variations with regard to article use because English articles are closely related to how a speaker conceptualizes nominal entities in a given discourse context. As such, speakers may impose a definite or indefinite status to a nominal depending on their perspective and interpretation of the referent in a particular context. Similarly, Butler's (1999) study recruited 20 native speakers of English who were experts in linguistics and English and asked them to take a fill-in-the-article test, which is similar to the one employed in the current study. In her study, the NES group did not score 100% either. The mean score of this group was 97.8% with a standard deviation of 3.23. Thus, we need to be mindful of individual variations in interpreting the scores. What is more meaningful in the think-aloud protocol is reasons that participants provided verbally for their article choices, which give us a better understanding of the underlying knowledge of articles. The reasons for article choices are analyzed and discussed in Section 5.2.3.

5.2.2. Summary of key findings: fill-in-the-article test.

Statistical analyses of the scores that the participants received on the fill-in-the-article test were reported in the above section. Key findings can be summarized as the following:

(1) Participants with higher levels of proficiency outperformed those with lower levels of proficiency.

(2) There was no statistically significant difference between the mean scores of the intermediate group and the advanced group.

(3) The result of the fill-in-the-article test indicates that the advanced group reached a near-native level with regard to article choices.

(4) The NES group did not score 100% in the fill-in-the-article test.

(5) Among the three choices available in the test, instances in which *a* was required were found to yield the highest scores in all learner groups.

(6) With regard to the variable of NP contexts for *the*, participants selected the form more accurately in anaphoric definite contexts than associative definite contexts, a tendency observed in all groups.

(7) Between the two categories of NP contexts for *a*, the beginners and intermediates performed better in referential indefinite contexts than non-referential contexts, but the advanced participants and NES participants performed with similar rates of accuracy in both contexts.

These quantitative outcomes produced tendencies and are illustrative of how the groups performed in the fill-in-the-article test, which is designed as a measure of explicit knowledge of articles. However, they do not inform us much about the underlying reasons behind such tendencies. Thus, it is essential to examine the verbal reports produced by the participants to understand what kinds of explicit knowledge they accessed when they were making article choices. In the next section, the underlying reasons for article choices are analyzed and discussed.

5.2.3. Reasons for article choices.

Following the report of how participants performed on the fill-in-the-article test, this section reports what kinds of reasons the participants provided for their article selections in the fill-in-the-article test. The verbal explanations for article choices were gathered through a think-aloud protocol. In the think-aloud task, the participants filled out 72 blanks with either *the*, *a*, or Ø (i.e., no article). As they were carrying out the task, they verbalized their thoughts and

provided reasons for their article choices. The analyses of the learner hypotheses about article use verbally expressed by the participants are discussed in this section to address the third research question that focused on the kinds of metalinguistic awareness Korean L2 speakers of English have at different proficiency levels and the fourth research question concerning how the learners' metalinguistic awareness may differ from that of native English speakers.

In analyzing explicit reasons for article choices, the following approach was applied. First, following Butler (2002), reasons for article selections were broadly divided into two: specific reasons (i.e., participants were able to identify rules of grammar or other reasons for their article selections) and non-specific reasons (i.e., they could not identify any specific reasons for their article choices). The non-specific reason group consisted of reasons based on any of the following three factors: plausible choices (e.g., "It sounds like it should be this article"), elimination (e.g., "Others don't work, so I think this article is the only choice, although I don't know why this is right") and no clue (e.g., I have no idea, so I just picked one")[10]. Accuracy with which articles were selected was examined for specific reasons as opposed to non-specific reasons.

Secondly, the specific reasons for article choices were further divided into six groups based on the kinds of distinctions that the participants made to select appropriate articles: (1) noun number (i.e., article selections are made based on whether a referent is singular or plural); (2) first vs. subsequent mention (i.e., article selections are made based on whether a nominal referent is mentioned for the first time or mentioned subsequently); (3) specificity (i.e., article selections are made based on whether a nominal refers to a specific referent or not); (4)

---

[10]   The examples for non-specific reasons are extracted from Butler (2002).

countability (i.e., article selections are made based on whether a nominal referent is countable or not); (5) combination (i.e., article selections are made based on a combination of the above distinctions); (6) non-generalizable reasons (i.e., article selections are made based on collocations or other idiosyncratic hypotheses). These groups were not established *a priori* but emerged as the think-aloud data were examined in a qualitative manner. After examining article reasons based on the above-mentioned groups, qualitative analyses of non-target-like article selections and reasons for them were conducted.

In the following sections, results of the analyses are presented by proficiency groups (i.e., the beginner group, the intermediate group, the advanced group, and the NES group). Then, a summary of key findings is provided.

(1) The Beginner Group

This segment reports the explicit reasons that the beginners provided for their article choices. First, the distribution of specific reasons and non-specific reasons in the beginner group is summarized in the following table.

Table 5. 9. Frequency of specific reasons and non-specific reasons in the beginner group

	N	Percentage
Specific reasons	565	78.4%
Non-specific reasons	155	21.6%
Total	720	100.0%

In this table, we observe that the beginners provided more specific reasons than non-specific reasons. Specific reasons were expressed for 78.4% of the article choices while 21.6% of the article choices were made without specific reasons. In other words, in 21.6% of the instances, beginners went for plausible choices, used elimination, or had no clue for their article choices. In

a similar study with Japanese L2 learners of English, Butler (2002) reported that learners at lower proficiency levels expressed specific reasons in 76.5% of the time, regardless of accuracy of article choices. Even though Butler's study did not analyze the data in the same manner as the current study, in general, both studies found that beginners tended to provide specific reasons more frequently than non-specific reasons.

Next, we examined whether specific reasons yielded more accurate article choices as opposed to non-specific reasons.

Table 5. 10. Accuracy of article choices with specific reasons and non-specific reasons

	Accurate article choice	Inaccurate article choice	Accuracy
Specific reasons ($N = 565$)	406	159	71.9%
Non-specific reasons ($N = 155$)	84	71	54.2%

$p < .0001$

The table shows that the majority of articles chosen with specific reasons were accurate. Specifically, when the beginners expressed specific reasons for their choices, 71.9% of them resulted in appropriate article selections. On the other hand, articles selected without specific reasons were accurate in only 54.2% of the contexts. As previously explained, non-specific reasons include situations in which articles were chosen based on plausible choices or guesses. With the beginners' interlanguage still in the process of development, some uses of articles in the fill-in-the-article text must have posed more challenges for these learners. When they were not able to provide specific reasons, the beginners may have not known why they chose the article or just guessed. In turn, when beginners explained article selections with specific reasons, their explanations were indicative of how they already developed hypotheses about article use in given contexts and they were also able to select articles accurately in the majority of the

instances in which they provided specific reasons for their selection.

Now we turn our attention to the kinds of specific reasons the beginners provided for accurate article selections. The first analysis is concerned with reasons for the selection of the definite article in the beginner group. Figure 5.1 illustrates the distribution of reasons that the beginners provided when they selected *the* accurately. As described in the introductionof this section, specific reasons were divided into six groups: (1) noun number, (2) first vs. subsequent mention, (3) specificity, (4) countability, (5) combination, and (6) non-generalizable reasons. Non-speficic reasons were also included in the figure to illustrate the percentage of the non-specific reasons in comparison to the specific reasons.

Figure 5. 1. Distribution of reasons provided by beginners when *the* was selected appropriately
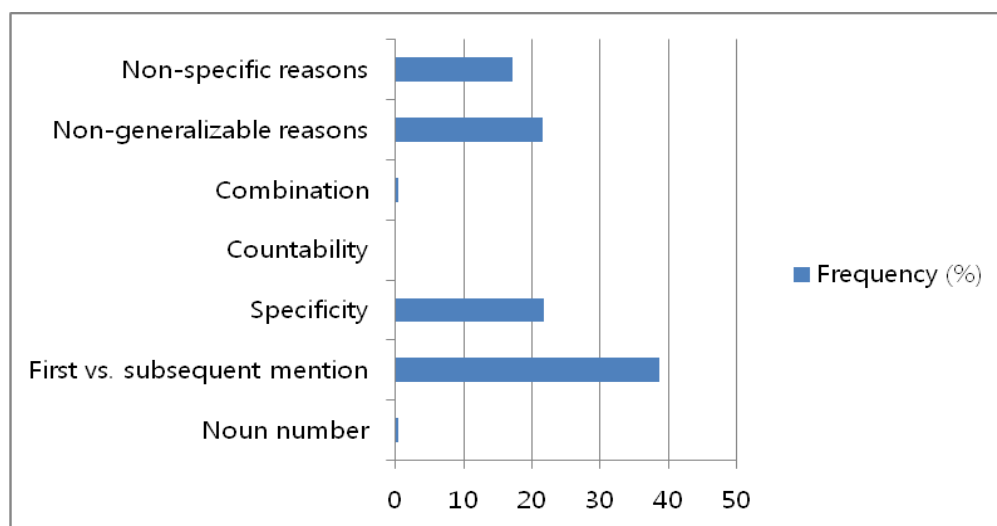
In this figure, we note that the most frequent reasons for the choice of *the* in the beginner group were the "first vs. subsequent mention" group. This group of reasons refers to explanations for article choices based on distinctions between instances in which referents were introduced for the first time and instances in which referents were mentioned subsequently after their initial introduction into the discourse. Therefore, this reason category is mainly applicable to 'anaphoric

definite' contexts for the definite article *the*. Examples of reasons provided by the beginners are in the following.[11] Underlined items are the focus of discussion.

(1) Answer: He filled (13) a pot with (14) water, put (15) the pot on (16) the
            stovetop, and turned on (17) the gas.
     Reason: It needs *the* because the pot is mentioned again. (B4)

(2) Answer: Oscar took (41) a cucumber out of (42) the fridge. He peeled it,
            sliced it, salted (43) the slices, and ate them all (…). Finished with
            (45) the cucumber, he sliced (46) a pear and ate it.
     Reason: The cucumber was mentioned before, so it should be *the*. (B5)

The above examples are illustrative of the beginners' ability to identify referents that were previously mentioned and appropriately select *the* when they were mentioned again. As observed in Table 5.8, the beginner group's accuracy in anaphoric definite contexts was considerably high at 76.7%. White (2010) reported that the use of *the* in subsequent mentions is one of the basic rules of articles covered in ESL textbooks. Thus, we can assume that the beginners in this study were aware of this rule and able to select *the* appropriately in anaphoric definite contexts.

Then, we note that the second most frequently cited factor was specificity of noun referents. In other words, beginners chose *the* when they perceived that a given nominal referred to a specific referent, as in the following examples:

(3) Answer: He filled (13) a pot with (14) water, put (15) the pot on (16) the
            stovetop, and turned on (17) the gas.
     Reason: I think it needs *the* because it refers to a specific stovetop. (B3)

(4) Answer: When (26) the water in (27) the pot began to boil, he grabbed (28) a
            box of (29) pasta from (30) the cabinet.
     Reason: I chose *the* because it is a specific cabinet. (B2)

---

[11] As explained in Chapter III, the participants were allowed to use Korean to provide verbal reports in the think-aloud protocol. Their comments were translated and summarized by the researcher, who is a native speaker of Korean.

Example (3) is concerned with the same item as example (1), but reasons provided were different. In example (1), B4 supplied *the* based on the recognition that *pot* was already introduced, while in example (3), B3 chose *the* based on the specificity of the referent. Different reasons provided for the same article choice are indicative of the value of think-aloud data as these qualitative data reveal underlying hypotheses that the learners apply in their article use.

The second highest reason supplied in the think-aloud was specificity. The frequent citation of 'specificity' as reasons for the choice of *the* suggests that beginners tend to associate the definite article with specificity of noun referents. The tendency to map *the* to the specificity of referents has been documented in the literature (Ionin et.al., 2004; Park, 2006; Thomas, 1989), and this tendency was evidenced in the beginners in this study. However, the reliance on noun specificity in the use of *the* may result in non-target-like use of *the*. In English, when a specific referent is introduced to the hearer for the first time, it is considered indefinite even though the referent is specific in the mind of the speaker. In other words, definiteness of a nominal is conditioned by the hearer's knowledge about the referent. As such, dependence on specificity of a noun referent may result in article errors like the following[12]:

(5) Answer: He filled (13) a pot with (14) water, put (15) the pot on (16) the stovetop, and turned on (17) the gas.
    Reason: I think it needs *the* here because it is a specific pot. (B6)
    the is needed here because it refers to a specific pot. (B2)

In the above example, *pot* was introduced into the discourse for the first time in item (13), which requires *a* as it is a singular entity. Yet, two beginners (B6 and B2) selected *the* based on the specificity of this referent. They failed to recognize that this referent is not yet known to the

---

[12] In this chapter, non-target-like article choices are marked with * to indicate inappropriate article use.

reader/hearer.

Finally, non-generalizable reasons were also cited frequently by beginners for the selection of *the*, including the following examples:

(6)   Answer: He filled (13) a pot with (14) water, put (15) the pot on (16) the
                  stovetop, and turned on (17) the gas.
      Reason: It is *the* because 'on the gas' is used together. (B1)

(7)   Answer: (47) The timer went off.
      Reason: You need *the* at the beginning of a sentence. (B4)

The above examples are illustrative of idiosyncratic hypotheses that beginners have developed with regard to the use of the definite article. Example (6) is an example of hypotheses based on collocations. Similarly, several beginners (B1, B3, B5) stated that *the* is typically used together with prepositions. B5 explained that *the* is used for expressions about locations. B6 used *the* for 'tools' such as *spoon* and *bowl*. Example (7) reflects an idiosyncratic rule. As Butler (2002) explains, L2 learners seem to go through a stage where they start to develop hypotheses that show sensitivity to wrong contexts and factors. This stage appears to begin when learners realize that textbook rules about articles may not apply to a variety of instances of article use (Pica, 1983). Then, they attempt to build hypotheses about article use based on collocations and other linguistic instances that they have been exposed to. However, non-generalizable reasons tend to diminish as proficiency improves further.

Up until this point, we examined reasons that the beginners provided when they selected the definite article accurately. The majority of article choices were made based on single factors (such as specificity and other reasons) when in fact they were supposed to consider various aspects of English articles and NP contexts. The beginners' tendency to select articles based on a single distinction is indicative of how the beginners' explicit knowledge of English articles is

incomplete and under constant revision and development.

Next, with regard to the accurate selection of *a*, a variety of explanations were expressed by the beginners. Figure 5.2 is illustrative of the distribution of reasons that the beginners provided when they selected *a* accurately.

Figure 5. 2. Distribution of reasons provided by beginners when *a* was selected appropriately
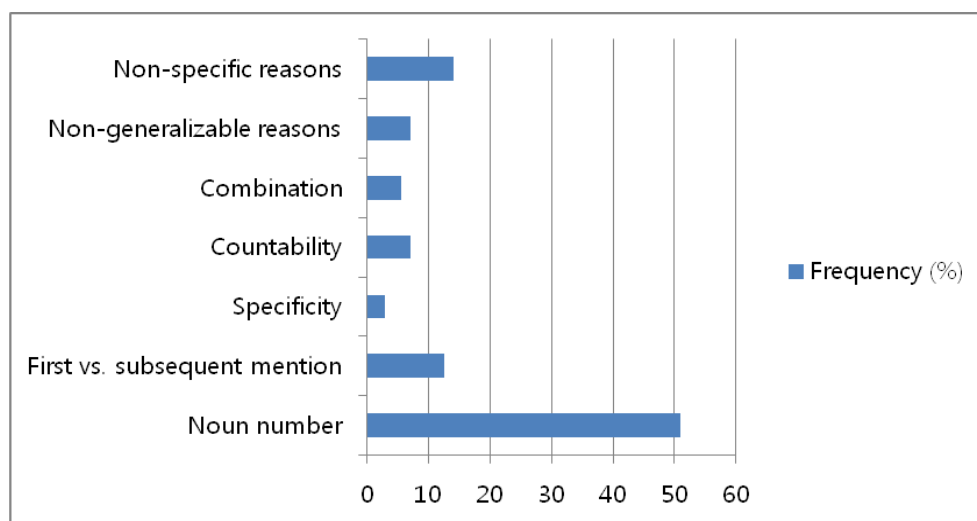
A key observation that emerged from this figure is that in the beginner group, the majority of the choices of the indefinite article were made based on noun number status of the referents. For example, the beginners accurately selected *a* in more than 50% of the instances based on the perception that noun referents were singular in number. The following examples are indicative of this tendency:

(8)   Answer: Oscar peeled and ate (12) a banana.
        Reason: You need *a* because it is one banana. (B3)

(9)   Answer: He chopped up part of (57) a red onion and mixed it into (58) the
                    pasta.
        Reason: I think it needs *a* here because it's one red onion. (B2)

In these examples, beginners selected *a* based on the perception that these nominal referents were singular entities. These reasons suggest that the beginners have an understanding of the

indefinite article as a marker of singularity of a nominal referent.

The second most frequently stated reasons produced by the beginner group for the choice of *a* were related to the distinction between first-mention and subsequent mention as in the following examples:

(10) Answer: Oscar peeled and ate (12) a banana.
     Reason: It's *a* because this is something new. (B4)

(11) Answer: Oscar poured (48) the hot water through (49) a plastic strainer.
     Reason: I use *a* because it is a new noun. (B5)

In the above examples, *a* was chosen because participants assumed that these nominals were mentioned for the first time. The use of *a* for the first mention and *the* for the subsequent mention is one of the basic rules of articles introduced in ESL textbooks, as reported by White (2010). Thus, it seems that beginners are aware of this rule and apply it appropriately.

Finally, the next figure illustrates the distribution of reasons that the beginners provided when they selected ∅ accurately.

Figure 5. 3. Distribution of reasons provided by beginners when ∅ was selected appropriately
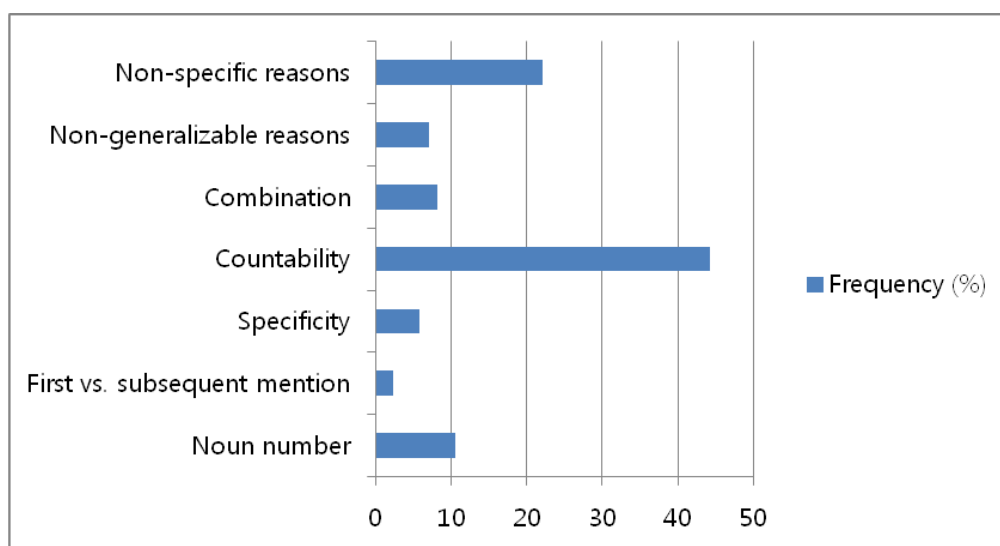
When beginners appropriately selected ∅ , the majority of the reasons they provided

were related to the countability of noun referents. In other words, they tended to choose Ø when noun referents were considered uncountable.

(12) Answer: He filled (13) a pot with (14) <u>water</u>, put (15) the pot on (16) the stovetop, and (…).
     Reason: You don't need any article here because *water* is uncountable. (B2)

(13) Answer: He added (52) a heaping tablespoon of (53) <u>butter</u> to (54) the hot pasta and (…).
     Reason: You cannot count *butter*, so you need no article. (B3)

In these examples, beginners perceived nouns such as *water* and *butter* as uncountable and selected no article accordingly. The use of no article for uncountable nouns is a primary rule that most ESL textbooks present, according to White (2010). In addition, nouns such as *water* and *butter* are often cited in these textbooks as examples of mass nouns. From the participants' comments, we thus can suggest that the leading factor that prompts the use of no article is countability.

While countability may be the most popular reason why no article was supplied in these contexts, the consideration of countability as the sole factor for the selection of Ø may result in article errors in other contexts, as illustrated in the following examples:

(14) Answer: Oscar poured (48) <u>the hot water</u> through (49) a plastic strainer, (…)
     Reason: You need *no article here because *water* is uncountable. (B3, B5, B10)

(15) Answer: He washed it down with (72) <u>the milk</u>.
     Reason: *I don't think you need an article here because *milk* is uncountable. (B5, B10)

(16) Answer: He chopped up part of (57) <u>a red onion</u> and mixed it into (58) the pasta.
     Reason: I remember *onion* is not countable, and *no article is needed here. (B5)

In examples (14) and (15), *water* and *milk* were perceived as uncountable and thus explained as not requiring an article. However, these beginners failed to notice that these referents were

already introduced into the discourse and were being mentioned subsequently, thus requiring the use of *the*. This observation suggests that they focused on the countability feature of these nominals without recognizing the entire discourse context. In example (16), participant B5 assumed that *onion* was uncountable and selected $\emptyset$ incorrectly. This error may be related to the misperception that countability is an inherent feature of nouns. Learners may be unaware that in English, countability of nominals is not a fixed attribute but dependent on contexts of usage. For instance, *chicken* can be countable when it refers to a bird, as in *The farmer has thirty chickens in his henhouse*, and it can be uncountable when it refers to meat, as in *I'm going to cook chicken for dinner*. Likewise, *onion* can be either countable or uncountable depending on which context this noun is used in. B5's analysis is exemplary of this perceived dichotomy, a fixed notion of countability of nouns.

The second most frequent reasons articulated by the participants for the choice of $\emptyset$ were non-generalizable reasons. The following utterances exemplify these instances:

(17) Answer: After that, he ate (18) a half cup of (19) assorted nuts.
     Reason: I think you need no article because there is a verb. (B3)

(18) Answer: He added two cups of (37) pasta to (38) the boiling water and set (39) the timer.
     Reason: I use no article when there is an amount. (B9)

(19) Answer: He sprinkled (62) garlic powder and grated (63) parmesan cheese onto (64) the pasta.
     Reason: I choose no article because this is about a state of some substance. (B6)

In example (17), B3 assumed that *assorted* was used as a main verb and the mistake in sentence parsing revealed an idiosyncratic hypothesis about article use. That is, this participant mistakenly analyzed *assorted* as a verb and determined that no article was needed for a verb. Similarly, hypotheses expressed by several beginners in examples (18) and (19) suggest that when they

cannot apply rules that they have learned, they tend to develop idiosyncratic hypotheses. For instance, B9 concluded that no article was needed for an expression of an amount, while B6 suggested that no article was used for an expression of a state of substance.

In sum, we have examined the hypotheses about English articles that beginners employed in making article choices. When beginners chose *the* accurately, these choices were attributed to the factors of subsequent mention and specificity of nominal referents. When *a* was selected accurately, the reasoning provided was mainly based on the singularity of referents. Finally, when beginners provided ∅ accurately, they did so primarily based on the perception that the concerned referents were uncountable.

We now turn to explore reasons behind the beginners' non-target-like article choices. Specifically, these instances of inaccurate article choices are examined according to four NP contexts: anaphoric definite, associative definite, referential indefinite, and non-referential indefinite contexts.

First, in anaphoric definite contexts, six beginners failed to choose the correct article for the following item.

(20)   Then he poured (55) the lemon juice onto (56) the pasta.

In the fill-in-the-article text[13], *lemon juice* had been previously introduced into the discourse when Oscar *sliced (23) a lemon in half, and then used (24) a squeezer to squeeze all the lemon juice into (25) an empty cup*. As the item numbers indicate, the first mention of *lemon juice* and its subsequent mention were separated by many sentences. Furthermore, the first mention of *lemon juice* and its second mention appeared in two separate pages. This distance between the first-

---

[13]   The entire text for the fill-in-the-article test is available in Appendix I.

mention and subsequent-mention may have caused difficulty for beginners to recognize the referential chain.

When participants did not recall that *lemon juice* was already introduced, they resorted to other features of the nominal. For instance, B1 and B4 chose *a*, which suggests that they perceived *lemon juice* as a countable noun. Specifically, B1 cited singularity of the referent as a reason for the choice of *a*, while B4 argued that it was a new referent. Other participants (B3, B6, B9, and B10) chose no article mainly because they perceived *lemon juice* as uncountable. These data suggest the compounding of multiple factors and the difficulties associated with accurate perception of countability.

As shown in Table 5.8 in the previous section, beginners were quite accurate in explaining *the* when it was used in anaphoric definite contexts (76.7%). Also, Figure 5.1 shows that the majority of beginners provided "first-mention vs. second mention" distinction as a basis for the selection of *the*. Therefore, errors committed for item (55) may not be a sign of the participants' inability to use *the* in subsequent mention contexts. Instead, the errors were possibly conditioned by the fact that the first mention was separated from the subsequent mention by too many sentences, which made it difficult for participants to recall the introduction of *lemon juice* into the discourse. In other words, the distance between the last mention of the referent or the use of the antecedent in a different expression (i.e., *lemon juice* vs. *juice*) may affect their explanations.

With regard to associative definite contexts, several beginners experienced difficulties in identifying the definite status of nominals that are considered definite by association with contextual features. For example, the following instances required *the*, but the majority of beginners failed to supply *the*.

(21) (…) he grabbed (28) a box of (29) pasta from (30) the cabinet.
(22) He added two cups of (37) pasta to (38) the boiling water and set (39) the timer.

In these associative definite contexts, *cabinet* and *timer* were considered definite even though they were not introduced in previous discourse. The setting of the story in the fill-in-the-article text is Oscar's kitchen. In this context, these nominals obtain definiteness based on cultural knowledge that suggests that certain pieces of furniture, appliances and other items would be present in a typical kitchen and thus uniquely identifiable in the mind of the hearer/reader. In this regard, Pica (1983) demonstrated that native speakers often use the definite article for first-time referents when they are identifiable in the immediate context of the discourse. However, this aspect of the articles seems to be a challenge to beginner-level learners. When beginners did not recognize these associative instances of the definite article, they made other choices and provided the following reasons.

(23) Answer: (…) he grabbed (28) a box of (29) pasta from (30) the cabinet.
Reasons: It's *a for a singular referent. (B3, B6)
It's *a for a new referent. (B4, B5)
It's *no article for an uncountable noun. (B7)
It sounds better with *no article. (B1, B8)

(24) Answer: He added two cups of (37) pasta to (38) the boiling water and set (39) the timer.
Reasons: It's *a for a singular referent. (B3, B9)
It's *a for a new referent. (B5, B7)
It's *a for a new and singular referent. (B4)
It sounds better with *a. (B1)

In example (23), seven beginners made incorrect article selections for *cabinet*. Four of them chose *a* based on either singularity or first-mention of the referent. The other three beginners selected ∅. In example (24), six beginners made inappropriate article choices for *timer* and expressed a variety of reasons. For L2 English learners from Korea at their beginner level of proficiency, their conceptualization of a prototypical kitchen may be different from that of native

213

English speakers. For instance, in Korean food culture, a timer is not used commonly for cooking. Therefore, we may carefully suggest that it might be a challenge for Korean participants to conjure up the image of a timer in a kitchen if they are not familiar with western style cooking.

Next, in referential indefinite contexts, one item was perceived to be difficult for beginners. As they made inaccurate article choices, the following reasons were provided.

(25) Answer: (…) and slowly ate half of (71) a five-ounce bar of Hershey's Special Dark candy.
 Reasons: It's *the for a specific bar. (B3, B5, B6)
 *The is used after a preposition. (B1, B7)
 It's *no article for an uncountable noun. (B10)

In this example, six beginners had difficulty with understanding the NP 'a bar of candy' and presented non-target-like hypotheses. Three beginners selected *the* and argued that it referred to a specific bar. One beginner posited a collocation-based rule which suggested that the definite article appears after a preposition. Another beginner perceived the candy bar NP as uncountable and thus selected ∅. A possible explanation for the difficulty with this NP is that the beginners may not be familiar with the use of *bar* as a measuring unit for nouns such as *candy*, *chocolate*, and *soap*. Further, an *ounce* is not a familiar unit of weight to Korean learners since they employ the metric system (i.e., gram).

Finally, in non-referential indefinite contexts, the following item caused inappropriate article choices among the beginners.

(26) Answer: (…) ate them all while he read This Week, (44) a popular weekly news magazine.
 Reasons: It is *the because the referent (*This Week*) was already mentioned. (B1, B8, B9)
 It refers to a specific magazine, so I chose *the*. (B2)
 *The is used before the word *magazine*. (B7)
 This NP describes what *This Week* is, so it needs *the*. (B6)

214

In the above example, six beginners failed to choose *a*. Instead, they selected *the* based on the perception that *magazine* refers to a particular referent that is unique and already known. This use of *a* is considered non-referential and it is employed to describe an attribute of a nominal referent, as in *Tom is a very smart boy*. Beginners in this study might have failed to identify non-referential uses of articles possibly due to the way this clause is constructed. In the above example, *a popular weekly news magazine* appeared following a comma in conjunction with *This Week*. It is possible that they would have selected *a* accurately if the clause was constructed as *This Week is a popular weekly news magazine*. Thus, the syntax in this utterance produced challenges for the beginners.

In this section, I dissected the explicit reasons provided by the beginners for their target-like and non-target-like article choices. The analyses of the think-aloud data revealed that most of the article choices were based on textbook rules and other simple rules about article use. It may be that they do not have a coherent and unified understanding of articles. When these rules did not help them, they resorted to non-generalizable reasons that were formulated, several of which were based on word collocations.

(2) The Intermediate Group

This section discusses metalinguistic understanding of English articles by the intermediate group based on the analyses of the explanations they supplied for their article choices. The organization of this section is similar to the beginner section. First, distribution of specific reasons and non-specific reasons and their rates of accuracy are examined. This segment is followed by a discussion of the reasons that the intermediates expressed for their selection of three choices (*the*, *a*, and ∅). Finally, reasons for inaccurate article choices are analyzed and

discussed to understand hypotheses about English articles behind non-target-like responses.

First, reasons provided by the intermediates were divided into two categories: specific reasons and non-specific reasons.

Table 5. 11. Frequency of specific reasons and non-specific reasons in the intermediate group

	N	Percentage
Specific reasons	633	87.9%
Non-specific reasons	87	12.1%
Total	720	100.0%

In this table, we find that specific reasons were provided much more frequently than non-specific reasons. For example, the intermediate speakers explained their article choices with specific reasons in 87.9% of the instances, while they did not express specific reasons in only 12.1% of the instances. In addition, the percentage of specific reasons in the intermediate group was higher than the beginner group (78.4%), an indication that intermediate participants have more meta-language for verbal explanations of grammatical features, in general, and metalinguistic knowledge of articles, in particular.

Next, accuracy for articles selected with specific reasons is compared with accuracy for articles selected with non-specific reasons.

Table 5. 12. Accuracy of article choices with specific reasons and non-specific reasons

	Accurate article choice	Inaccurate article choice	Accuracy
Specific reasons ($N = 633$)	513	120	81.0%
Non-specific reasons ($N = 87$)	55	32	63.2%

$p < .0001$

The table shows that the majority of articles chosen with specific reasons were accurate. When specific reasons were provided, 81% of them resulted in appropriate article selections. In turn, accuracy in which articles chosen without specific reasons was 63.2%. For both cases, rates

of accuracy (81.0% for specific reasons and 63.2% for non-specific reasons) were higher than those in the beginner group (71.9% for specific reasons and 54.2% for non-specific reasons). These findings suggest that intermediate participants have more metalingual expressions to explain their article choices and more accurate knowledge of English articles compared to the beginners.

With descriptive statistics about article choices and reasons for the selection of article forms provided by the participants in the fill-in-the-article test, it is necessary to investigate specific reasons that the intermediates verbally reported when they were selecting articles in the fill-in-the-article test. Figure 5.4 illustrates frequencies of reasons that intermediate speakers employed when they selected *the* accurately.

Figure 5. 4. Distribution of reasons provided by intermediates when *the* was selected appropriately
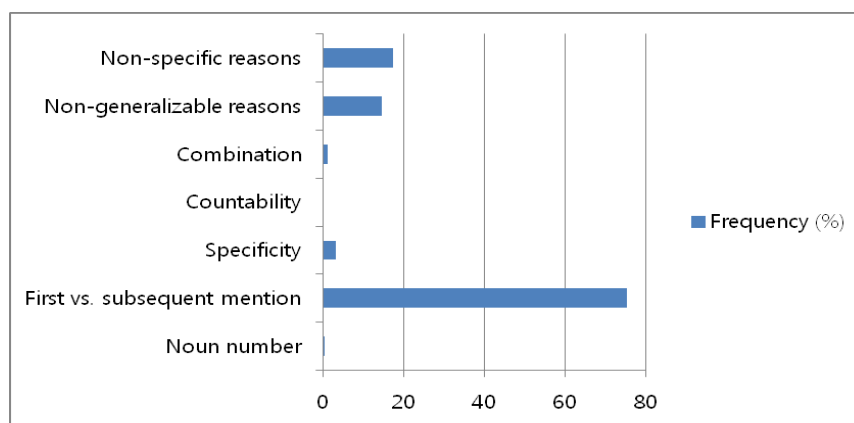
In Figure 5.4, we note that the most frequent reasons that intermediate speakers expressed for their choice of *the* were related to the distinction of first vs. subsequent mention. More than 70% of the instances in which *the* was accurately selected were based on the participants' awareness that a noun referent had been previously introduced and was mentioned subsequently. The distinction between first-mention and subsequent-mention is applicable to

anaphoric definite contexts that require the use of *the*. This observation may explain the high rate

of accuracy (93.9%) with which intermediates used *the* in anaphoric definite contexts. In other

words, intermediate speakers seem to be able to identify anaphoric relationships of noun

referents in a coherent discourse and accurately provide the definite article. Examples of such

instances are provided in the following:

(27) Answer: He filled (13) a pot with (14) water, put (15) the pot on (16) the stovetop,
                and…
      Reason: It needs *the* because the pot is mentioned again. (I1)

(28) Answer: When (26) the water in the pot began to boil, he grabbed (28) a box of
                (29) pasta from (30) the cabinet.
      Reason: *water* was mentioned before, so it should be *the*. (I7)

In both examples, intermediate participants were quick to point out that these nominals were

already introduced in the discourse, thus taking *the* when they were mentioned again. The

intermediate group exhibited 93.9% accuracy in the selection of *the* for anaphoric expressions.

The next most frequent reasons were non-generalizable reasons and non-specific reasons,

although their frequencies were far lower than the most frequent reasons. We now turn to

examine several of the non-generalizable reasons expressed by the intermediate group with

regard to the use of the definite article.

(29) Answer: (…) put (15) the pot on (16) the stovetop, and turned on (17) the gas
      Reason: *the* is needed because it is a particular location. (I7)

(30) Answer: On (34) the side, in (35) small print at (36) the very bottom, was (…)
      Reason: *the* is used for a specific location. (I3, I4, I7, I8)

In examples (29) and (30), we note that several intermediate participants have a non-

generalizable hypothesis that *the* is used for expressions related to locations. One possible

explanation for this hypothesis is that location-related expressions are often used in instances in

which referents are visible or identifiable in the immediate environment of the discourse. Thus,

constructions that refer to particular locations tend to require the use of *the*. As such,

intermediates may draw a generalization from similar instances that they have been exposed to

and concluded that *the* is always used in nominals that refer to particular locations.

Other non-generalizable reasons were also provided for the use of *the*.

(31) Answer: (…) (1) the less people eat, (2) the longer they live.
     Reason: it's a fixed expression, as in '*the* + comparative, *the* + comparative'.
     (I5, I7, I9)

(32) Answer: He loved (20) the pecans best, although they were (21) the fewest in
     (22) the assortment.
     Reason: *the* is used before a superlative form of an adjective. (I1, I2, I3, I4,
     I5, I6, I7, I8, I9, I10)

The above examples are illustrative of other non-generalizable hypotheses expressed by

intermediate participants. Here, they explained the choice of *the* with specific article rules that

they have probably learned from textbooks or during instruction. In a review of several ESL

textbooks which explain article use, White (2010) found that many of the criteria presented in

these ESL grammar books for making article choices are based more on collocations, particular

words or constructions rather than conceptualizations that a speaker wishes to convey or broader

discursive contexts. The hypotheses provided in the above examples are part of these

collocation-based article rules. As White (2010) argued, these fragmented and collocation-based

rules may do disservice to ESL learners because such article rules do not explicitly state why

certain nominals are consider definite and others are not.

Next, with regard to the indefinite article, a variety of reasons were stated by the

intermediates. The accuracy with which the intermediate speakers seletected *a* was very high

(92.0%). Figure 5.5 shows the distribution of reasons that the intermediates provided when they

supplied *a* accurately.

Figure 5. 5. Distribution of reasons provided by intermediates when *a* was selected appropriately
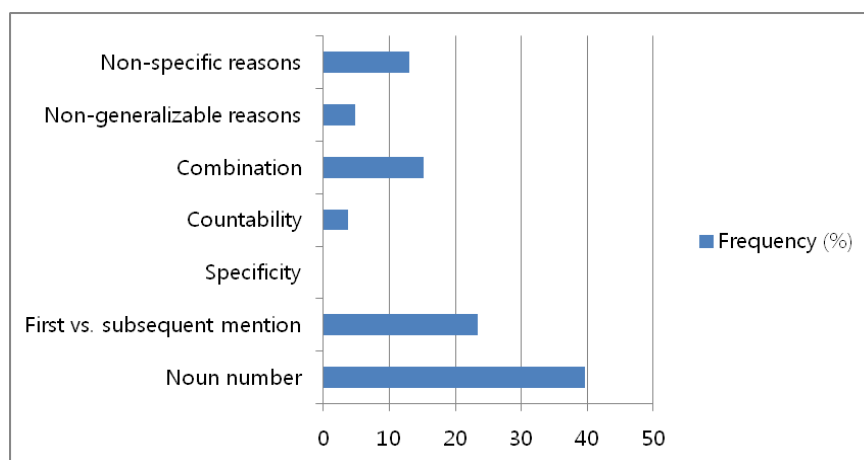
In this figure, we find that a wide range of reasons were supplied by the intermediates. In fact, the distribution of reasons was much wider than those we noted for the uses of *the*, which was mainly chosen based on subsequent-mention. In general, we find that more factors converge to determine the uses of *a*. The selection of *a* implies that a nominal is countable, singular, and not identifiable in the mind of the hearer. Thus, features such as countability, noun number, and first-mention contribute to the target-like selection of the indefinite article. In the following, we examine which features were considered most frequently by the intermediate group.

The most frequent reasons provided by the intermediate group for the selection of *a* were based on the factor of noun number. That is, nearly 40% of the instances in which *a* was accurately supplied were based on the perception that a noun referent was singular, as in the following examples:

(33)   Answer: (…) then used (24) a squeezer to squeeze all the lemon juice into (…)
           Reason: It is *a* because he used one squeezer. (I1)

(34)   Answer: He put (65) a spoon into (66) the bowl, sat down at his dining table,
                          and dug in.
           Reason: It is one spoon, so it needs *a*. (I5)

In these examples, the indefinite article appeared in referential indefinite contexts, in which a

specific yet unknown to the hearer is mentioned for the first time. The accuracy for *a* in these

contexts was very high (95.3%) in the intermediate group, and in these instances, intermediates

tended to focus more on the feature of singularity of nominal referents. As White (2010)

maintains, *a* is an article for individuation. Therefore, it denotes a singular countable referent that

is conceptualized in the minds of the speaker and the hearer as an entity with a clear boundary.

The frequent citation of singularity as a reason for the choice of *a* may suggest that intermediates

have an explicit understanding about this aspect of the indefinite article.

The second most frequently stated reasons in the intermediate group for the choice of *a*

were concerned with the distinction between first-mention and subsequent-mention.

(35) Answer: Oscar poured (48) the hot water through (49) a plastic strainer (...)
      Reason: It is *a* because it is a new noun. (I7)

(36) Answer: He put (65) a spoon into (66) the bowl, sat down at his dining table,
               and dug in.
      Reason: This is a new thing, so it needs *a*. (I9)

These examples also belong to referential indefinite contexts, and the first-mention feature was

cited by some intermediates as a reason for the selection of *a*. It is worth noting that the same

participants seem to fluctuate between the features of singularity and first-mention. In other

words, the same participant would cite singularity of a nominal as a reason for the choice of *a* in

some instances, yet in other similar instances, first-mention was cited as a reason. These

fluctuations seem to point to the fact that these intermediate participants indeed considered both

aspects of indefinite nominals but possibly verbalized only one of the two aspects as a primary

factor for their article selections.

While the majority of intermediates expressed either singularity or first-mention as

reasons for the choice of *a*, Figure 5.5 also illustrates that the combination of factors was

relatively frequent in the intermediate group as shown in the following examples:

(37) Answer: He sliced (23) a lemon in half, and then used (24) a squeezer (…)
      Reason: It is *a* because *lemon* is new and it's singular (I10)

(38) Answer: (…) and slowly ate half of (71) a five-ounce bar of Hershey's Special
                 Dark candy.
      Reason: This is a new thing and singular, so I used *a* (I4)

In these examples, a combination of different features were considered when intermediates

selected *a*. The most common combinations were singularity and first-mention, but a

combination of singularity, countability and first-mention were also cited as reasons for the

choice of *a*. These findings suggest that participants in the intermediate group are conscious of

various distinctions with regard to indefinite singular nominals.

Finally, with regard to the accurate selection of ∅ , the following figure illustrates the

reason categories.

Figure 5. 6. Distribution of reasons provided by intermediates when ∅ was selected appropriately
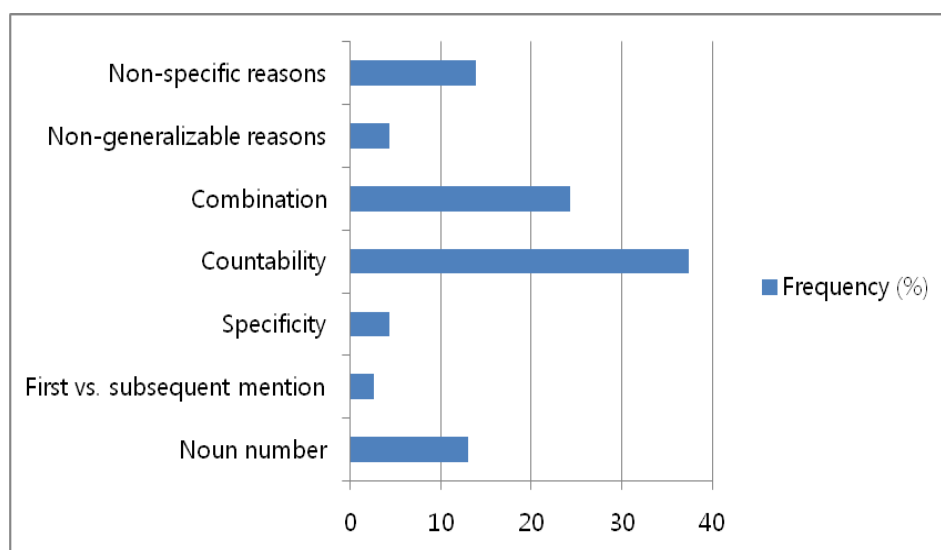
In the above figure, we find that the most frequently cited factor was countability,

followed by the combination and the noun number. We first examine examples in which

countability was provided as a main reason for the selection of ∅ :

(39) Answer: He filled (13) a pot with (14) water (…)
     Reason: You don't need any article because *water* is not countable.(I6)

(40) Answer: He grabbed (28) a box of (29) pasta from (30) the cabinet.
     Reason: No article is needed because *pasta* is uncountable. (I5)

According to these examples, intermediates seem to have a hypothesis which suggests that an

article is not required if a referent is uncountable. As discussed in detail in Chapter II, *the* is

required when a nominal referent is definite, regardless of its countability or number. When a

nominal referent is not definite but is in need of conceptualization as a singular entity with a

clear boundary, *a* is required. In all other instances, an overt article is not required when

indefinite referents are uncountable or plural. Thus, the think-aloud data suggest that

intermediates consider countability as a key factor in the selection of ∅ .

The second most frequent reasons were combinations, which means that more than one

feature of a nominal was considered, as in the following examples:

(41) Answer: He added (52) a heaping tablespoon of (53) butter to (54) the hot pasta
     Reason: It is uncountable and not mentioned before. (I5)

(42) Answer: These studies are based on (3) experiments with (4) mice and other
            small animals
     Reason: I chose no article here because *mice* is plural and they are not
            particular mice but general mice. (I8)

In example (41), the participant considered first-mention and countability as factors for the

selection of ∅ , while in example (42), plurality and non-specificity as reasons for ∅ . As

discussed in Figure 5.5, intermediates tend to consider more features of nominals in article

choices compared to the beginner group. Again, these findings seem to point to the fact that

intermediates have a deeper and more comprehensive understanding of the variable uses of

articles in several contexts.

The third most frequently cited factor was noun number, which is related to the use of Ø when indefinite referents are plural. Let us examine the following examples:

(43) Answer: These studies are based on (3) experiments with (4) mice and other small animals

Reason: *Mice* is plural, so you don't need any article. (I6)

(44) Answer: After that, he ate (18) a half cup of (19) assorted nuts.

Reason: I picked no article because it is plural. (I7)

In these examples, plurality of nominals was cited as a primary reason for their choices of no article. In addition, the fact that they did not choose *the* for these NPs suggests that the intermediates already identified them as indefinite nominals implicitly but did not explicitly state reasons for why they were considered indefinite. Instead, these participants focused on the plurality as the most prominent reason in these instances.

In summary, the intermediate participants cited subsequent-mention of nominals as a primary reason for the choice of the definite article. The selection of *a* was mainly explained in terms of singularity and first-mention of noun referents. In addition, Ø was chosen based on factors such as non-countability, combination of countability, noun number and first-mention, or plurality of nominals. As seen earlier in Table 5.7, intermediates as a group outperformed the beginners on the fill-in-the-article test. In particular, they scored the highest with regard to the selection of *a* among the three groups. If we couple this information with the performance results on the fill-in-the-article test, the metalinguistic hypotheses expressed in the think-aloud protocol suggest that intermediate participants in this study have sensitivity towards various aspects of English articles and possess conscious knowledge of how articles should be used.

We now turn our attention to reasons for non-target-like article selections in the

intermediate group. Similar to the analyses in the segment for the beginner group, several

instances in which intermediates chose inappropriate articles are examined qualitatively to

determine the aspects of English articles that posed difficulties and examine the metalinguistic

hypotheses they employed to cope with these difficulties.

     Recall that in anaphoric definite contexts, the intermediate group performed with a high

degree of accuracy. As shown in Table 5.8, the accuracy for *the* in anaphoric definite contexts

was 93.9%. However, the group's performance in associative definite contexts was much lower

with the accuracy for *the* in these contexts at only 52.6%. The following items were found to be

challenging to most intermediate participants.

(45)   He filled (13) a pot with (14) water, put (15) the pot on (16) the stovetop, and
          turned on (17) the gas.
(46)   (…) he grabbed (28) a box of (29) pasta from (30) the cabinet
(47)   He added two cups of (37) pasta to (38) the boiling water and set (39) the timer.
(48)   Oscar took (41) a cucumber out of (42) the fridge.

Nominals in these examples refer to kitchen-related referents. Since the setting of the story is

Oscar's kitchen, these NPs are considered definite in English by association with a typical

western kitchen that interlocutors picture in their minds. In the context of the story, the narrator

assumes that these referents would be perceived as unique in Oscar's kitchen in the mind of the

reader. However, many intermediate participants in this study failed to recognize the definite

status of these nominals in this particular discursive context. Instead, they chose *a* or *Ø* and

provided a variety of reasons. In the following, reasons cited by intermediate participants are

examined for each item:

(49)   Answer: He filled (13) a pot with (14) water, put (15) the pot on (16) the stovetop,
                and turned on (17) the gas.
       Reasons: It is *a* because it was not previously mentioned. (I3, I4, I5)
                I'm not sure, but I think it's *a*. (I1, I2)
                It is *a* for a countable noun. (I9)

It is *a* for a singular noun. (I8)
It is *a* for a countable, singular noun not previously mentioned. (I10)
I'm not sure, but I think it needs *no article. (I7)

In the above example, nine intermediates out of ten failed to choose *the*. The majority of them selected *a* based on the factors of first-mention, singularity, or combination of the two. One participant selected ∅ but was not able to provide a specific reason. Here, we note that the intermediates failed to recognize the definite status of *stovetop* in association with *kitchen*. Thus, they turned to singularity and other aspects of the nominal to make an article choice.

For item (17), intermediates provided the following reasons:

(50)   Answer: He filled (13) a pot with (14) water, put (15) the pot on (16) the stovetop, and turned on (17) the gas.
       Reasons: It's * no article for an uncountable noun. (I3, I4, I5, I6, I8, I9, I10)
                 It's * no article for an uncountable noun that was not mentioned before. (I1)

In this instance, eight intermediates selected ∅ instead of *the*, and most of them based their decisions on the perception that *gas* is uncountable. The tendency to consider *gas* as uncountable may be attributed to how most ESL textbooks teach noun countability. As White (2010) describes, most textbooks seem to treat countability as inherent attributes of certain noun categories, such as mass nouns and abstract nouns. However, in English, the same nouns may be conceptualized differently and thus take on different articles depending on the context in which they appear. In the above example, turning on *the gas* refers to the action that allows gas to be supplied to the stove for cooking. In this kind of context, *gas* is used with the definite article. Most intermediates do not seem to have a grasp of this kind of article usage.

The following are reasons that intermediates provided for item (30):

(51)   Answer: (…) he grabbed (28) a box of (29) pasta from (30) the cabinet.
       Reasons: It's *a* for a new noun. (I4, I5, I7, I10)

It's *a* for a singular noun. (I1, I3, I6)
It's *a* for a countable noun. (I8, I9)
I'm not sure, but I think it's *no article (I2)

In this instance, all the intermediates failed to select *the*. Instead, the majority of intermediates

chose *a* based on a variety of factors such as first-mention, singularity, and countability. One

intermediate was not sure but selected Ø without a specific reason.

Reasons for item (39) were similar to those presented in the above example:

(52) Answer: He added two cups of (37) pasta to (38) the boiling water and set (39)
                the timer.
      Reasons: It's *a* for a new noun. (I4, I5, I6, I7, I10)
                It's *a* for a new, singular noun. (I3)
                It's *a* for a countable noun. (I9)
                It's *a* for a new, countable noun. (I8)
                I'm not sure, but I think it's *a*. (I2)

All the intermediate participants selected *a* instead of *the* and cited first-mention, singularity, and

countability as primary reasons for their article choices. It seems that these participants did not

recognize *timer* as a uniquely identifiable item in the kitchen and thus failed to perceive it as a

definite nominal.

In a similar vein, item (42) caused confusion among the intermediate participants.

(53) Answer: Oscar took (41) a cucumber out of (42) the fridge.
      Reasons: It's *a* for a new noun. (I4, I10)
                It's *a* for a new, singular noun. (I3, I5)
                It's *a* for a countable noun. (I9)
                I'm not sure, but I think it's *a*. (I1, I8)
                I'm not sure, but I think it's *no article. (I2, I6)

In this example, nine intermediates failed to select *the*. Instead, *a* was chosen on the basis of such

factors as first-mention, singularity, countability, and guessing. Two intermediates guessed that it

would be an instance for Ø. As one expects to find a fridge in a typical kitchen, the fridge in

Oscar's kitchen is considered definite in the context of the story. However, the majority of

intermediates failed to recognize the uniquely identifiable status of *fridge* in Oscar's kitchen. Only one intermediate (I7) correctly selected *the* and stated, "I learned in my English class that I need to use *the* for *fridge* because I made a mistake of saying '*a fridge*'."

In the above examples for items (16), (17), (30), (39), and (42), these nominal referents acquire a definite status by association with *kitchen* and based on shared cultural knowledge of what constitutes a typical kitchen. This particular use of *the* was classified by Liu and Gleason (2002) as cultural use of the definite article. The researchers explored whether different uses of the definite article present different levels of difficulties for ESL learners and reported that the cultural use seems to be acquired last. In this current study, the intermediates performed with high scores in anaphoric definite contexts, but their performance was much lower in associative definite contexts, an observation which lends support to Liu and Gleason's (2002) findings.

With regard to NP contexts that require the indefinite article, the intermediate participants scored high in referential indefinite contexts with the accuracy of 95.3%. The accuracy was lower in non-referential indefinite contexts (82.0%). One particular instance that most intermediate participants made non-target-like choices is the following and they provided a variety of reasons:

(54)   Answer: (…) ate them all while he read This Week, (44) a popular weekly news magazine.
        Reasons: It's *the* for the reference to a particular magazine. (I3, I5, I6, I9)
                It's *the* because This Week is a magazine that people know. (I1)
                It's *the* for a unique noun. (I8)
                It's *no article for an uncountable noun. (I7)

This item also posed difficulty to the beginners, and it was also perceived difficult by the intermediate group as seven intermediates failed to choose *a*. The majority of the intermediate participants selected *the* based on the perception that *magazine* refers to a particular referent that

is unique and already known. One intermediate selected ∅ because *magazine* was considered uncountable.

As discussed earlier in the segment for the beginner group, *a* is used in this type of non-referential contexts to describe an attribute of a nominal referent. Similar to the case of the beginner group, intermediates might have failed to recognize this particular use of *a* possibly because of the unfamiliar syntactic pattern of this sentence.

This segment examined various metalinguistic hypotheses about English articles that the intermediate participants provided for their target-like and non-target-like article choices in the fill-in-the-article test. In general, the intermediate group seems to have a more comprehensive understanding of English articles compared to the beginner group and possess sufficient metalanguage to explicitly explain reasons behind their article choices. However, they were found to have difficulties with the use of *the* in associative definite contexts when *the* was required for nominal referents deemed definite by cultural associations.

Continuing with the analyses to address the third research question that focuses on the kinds of metalinguistic knowledge of articles according to proficiency levels, the next segment reports the data from the think-aloud protocol provided by the advanced group.

(3) The Advanced Group

In this section, results of the think-aloud protocol of the advanced group are reported. First, reasons provided by the advanced speakers were divided into specific and non-specific reasons, which are summarized in the following table.

Table 5. 13. Frequency of specific reasons and non-specific reasons in the advanced group

	N	Percentage
Specific reasons	523	72.6%
Non-specific reasons	197	27.4%
Total	720	100.0%

In this table, we find that advanced speakers tended to express specific reasons more frequently than non-specific reasons. Specific reasons accounted for 72.6% of the instances while no specific reasons were provided in 27.4% of the cases. When compared with the intermediate group, the percentage of non-specific reasons in the advanced group is higher than that in the intermediate group (12.1%). At the advanced level of proficiency, non-specific reasons may suggest that those article choices were made with implicit knowledge, which could not be explicitly explained by advanced speakers.

The next table compares the accuracy in which articles were chosen when participants provided specific reasons and when non-specific reasons were expressed.

Table 5. 14. Accuracy of article choices with specific reasons and non-specific reasons

	Accurate article choice	Inaccurate article choice	Accuracy
Specific reasons ($N = 523$)	421	102	80.5%
Non-specific reasons ($N = 197$)	156	41	79.2%

$p < .0001$

According to the table, article choices made by advanced speakers were highly accurate in both instances in which article choices were accompanied by specific reasons and non-specific reasons. When advanced speakers provided specific reasons, those article choices were accurate in 80.5% of the instances. In turn, the accuracy with which articles chosen without particular reasons was similarly high, 79.2%.

It is important to note that the accuracy for articles selected without specific reasons was much higher in the advanced group (79.2%) than the intermediate group (63.2%) and the beginner group (54.2%). In other words, the use of intuition by advanced participants resulted in more accurate article choices, an indication of more target-like implicit knowledge exhibited in the advanced group.

Next, we draw our attention to the kinds of specific reasons that advanced speakers provided for each target article. Qualitative analyses help us develop a better understanding of the kinds of metalinguistic hypotheses the advanced speakers employed while they were supplying articles in the fill-in-the-article test. First of all, we examine the distribution of reason categories that the advanced speakers provided when they selected *the* accurately.

Figure 5. 7. Distribution of reasons provided by advanced speakers when *the* was selected appropriately
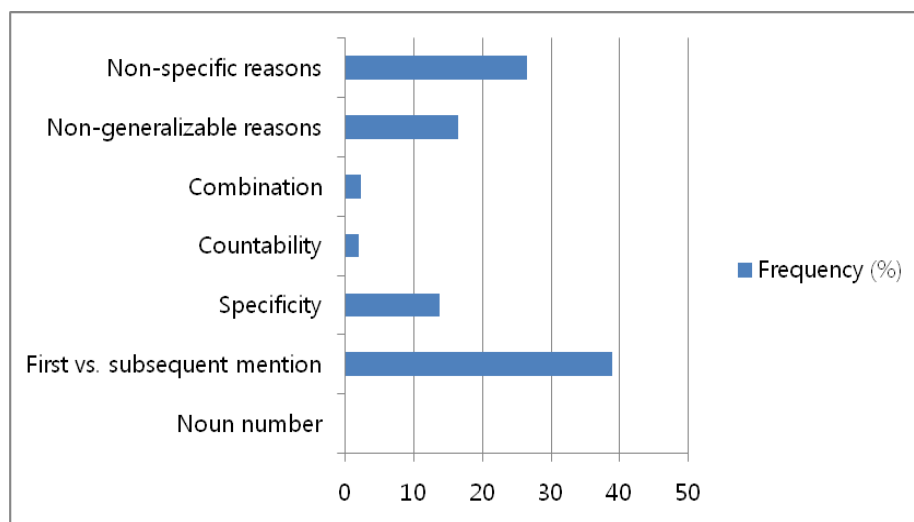
This figure shows that the most frequent reasoning that advanced speakers used in the accurate choices of the definite article was the distinction of first vs. subsequent mention. When a previously introduced referent was mentioned again (i.e., in anaphoric definite contexts), advanced participants selected *the*, as in the following examples:

(55) Answer: When (26) the water in (27) <u>the pot</u> began to boil, he grabbed (…)
       Reason: You need *the* here because the pot was introduced before. (A2)

(56) Answer: He loved (20) the pecans best, although they were (21) the fewest in
                     (22) <u>the assortment</u>.
       Reason: It is *the* because it was mentioned before as *assorted nuts*. (A4)

In these examples, advanced participants accurately identified the anaphoric relationships of noun referents and accurately selected the definite article. As discussed in Table 5.8, accuracy in the use of *the* in anaphoric definite contexts was 86.1% in the advanced group. Thus, the data suggest that advanced speakers are cognizant of the use of *the* for anaphoric expressions.

The second most frequent reasons that emerged from the data were the non-specific reasons. That is, they made article choices based on reasons such as "sounds right" and "feels right". In these instances, article choices were made accurately, which suggests that advanced speakers resorted to their implicit knowledge of English articles. The majority of the instances in which *the* was chosen without specific reasons belonged to associative definite contexts. In short, when *the* was required in anaphoric definite contexts, advanced participants were aware of the subsequent-mention and selected *the* correctly, as discussed in the above examples. In turn, when the definite article was required in associative definite contexts, the participants mainly resorted to strategies such as 'feel right' and 'sound right' and selected *the*.

The third most frequent reasons were non-generalizable reasons. Examples of non-generalizable reasons are in the following:

(57) Answer: (…) (1) <u>the less</u> people eat, (2) <u>the longer</u> they live.
       Reason: It is a fixed expression (*the* + comparative, *the* + comparative). (A1)

(58) Answer: Oscar took (41) a cucumber out of (42) <u>the fridge</u>.
       Reason: You use *the* for locations. (A3)
                 The comes after 'of'. (A7)
                 Fridge takes *the*. (A5)

(59)   Answer: On (34) the side, in (35) small print at (36) <u>the very bottom</u>, was (…)
        Reason: *The* is used before 'very'. (A8)

These excerpts show that several advanced participants employed non-generalizable hypotheses

in the selection of *the* that are mainly based on collocations.

Next, the reason categories for the accurate selection of the indefinite article were

examined. The following figure shows the distribution of reasons that the advanced group

provided when they selected *a* accurately.

Figure 5. 8. Distribution of reasons provided by advanced speakers when *a* was selected
appropriately
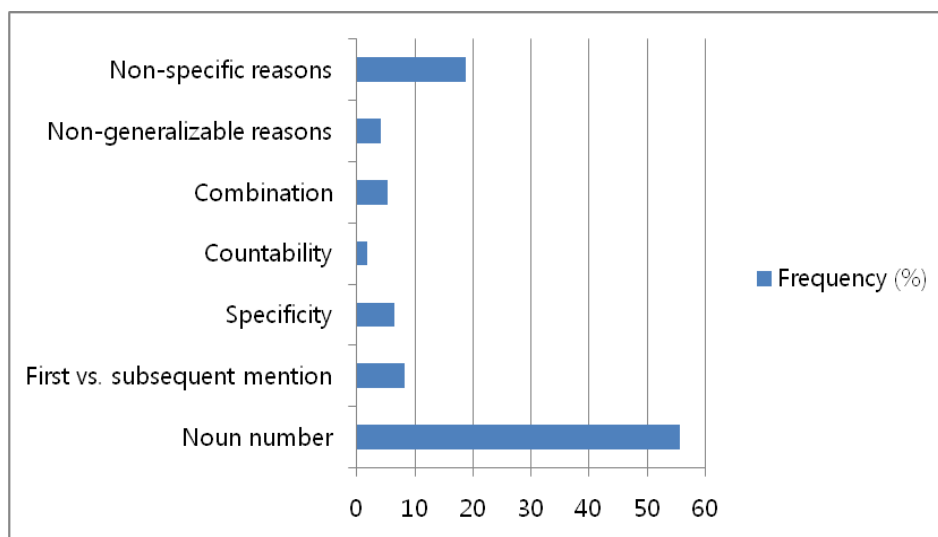
In the advanced group, the most frequently cited factor for the use of *a* was noun number.

That is, a majority of advanced participants cited singularity of noun referents as a primary

reason for their choices of *a*, as in the following examples:

(60)   Answer: (…) and then put (50) the pasta into (51) <u>a bowl</u>.
        Reason: You need *a* because it is just one bowl. (A4)

(61)   Answer: Then he poured (69) cold nonfat milk into (70) <u>a large glass</u> (…)
        Reason: It is *a* for a singular thing. (A6)

In the above examples, the participants identified that respective noun referents were singular

and supplied *a*. The role of *a* was defined by Christophersen (1939) as a denotation of unity, which turns a continuate-word into a unit word. In other words, the appearance of *a* indicates that a referent is a unit entity in its singular form. Thus, the reasons verbalized by the advanced speakers in the above examples suggest that they have an understanding that the primary semantic property of *a* is of singularity.

The second most frequent reasons employed by the advanced participants for the choice of *a* were non-specific reasons. As discussed earlier for the definite article, non-specific reasons in the advanced group indicate that the participants utilized their implicit knowledge of articles which could not be explicitly verbalized in a conscious manner. In other words, they could not explain why, but they knew when to use *a* and selected the article accurately based on intuitive judgments.

Finally, the following figure examines the reasons that the advanced speakers provided when they selected $\emptyset$ accurately.

Figure 5. 9. Distribution of reasons provided by advanced speakers when $\emptyset$ was selected appropriately
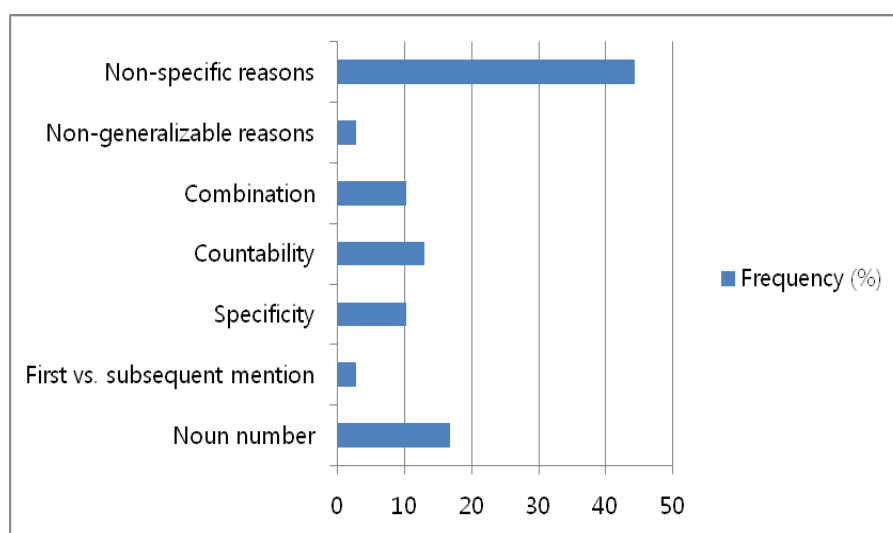
In this figure, we note that the most frequently cited reasons were non-specific reasons. In other words, when advanced speakers selected the option of ∅, they did not provide particular reasons behind such choices but depended on their intuition of what sounded right. In instances in which an article was not required, non-specific reasons were cited much more frequently than other factors such as noun number, countability, specificity, and combination. This tendency relying on non-specific reasons is indicative of the participants' implicit knowledge of articles in relation to their article choices.

The next most frequently cited factor for the accurate selection of ∅ was noun number, and this factor was mostly relevant to instances in which plural nouns were used, as in the following examples:

(62) Answer: These studies are based on (3) experiments with (4) mice and other small animals
     Reason: You don't need anything because it is plural (A3)

(63) Answer: On (32) the front, in (33) large letters, was "Barilla Penne – Italy's #1 Brand of Pasta".
     Reason: It does not need any article for a plural noun. (A8)

In the above examples, we find that advanced participants knew that an article is not required for plural referents in indefinite contexts and made appropriate article choices.

In short, target-like selections of *the* were made primarily based on subsequent-mention of nominal referents and intuitive judgments. When advanced participants chose *a* appropriately, they considered singularity of referents as a main factor. In turn, when ∅ was selected appropriately, such decisions were mostly based on intuitive judgments. In the advanced group, therefore, a tendency to rely on intuitive feel was observed for target-like article choices. This observation suggests that much of their understanding of English articles has become implicit knowledge as their proficiency level has improved.

235

So far, the reasons for target-like article choices were discussed. We now turn to reasons for non-target-like article selections by the advanced group. In the following, instances in which the majority of advanced speakers supplied inappropriate articles are examined according to NP contexts.

First, in anaphoric definite contexts, the advanced participants performed satisfactorily, achieving an accuracy of 86.1%. In associative definite contexts, six advanced participants failed to select *the* for the following item:

> (64)   Answer: He added two cups of (37) pasta to (38) the boiling water and set (39) the timer.
>
> Reasons: It's *a that describes *timer*. (A2)
> It's *a for a non-specific, singular item. (A4)
> It's *a for a singular noun. (A6, A7)
> It's *a for the first-mention. (A9)
> I'm not sure, but I think it's *a. (A5)

These participants selected *a* based on perceptions that *timer* is singular, non-specific or mentioned for the first time. This instance was also difficult for both beginners and intermediates.

In referential indefinite contexts, one particular instance caused six advanced participants to select *the* instead of *a*.

> (65)   Answer: (…) sat down, and slowly ate half of (71) a five-ounce bar of Hershey's Special Dark candy.
>
> Reasons: It's *the for a specific amount. (A3)
> It's *the because adjective makes it specific. (A5)
> It's *the for specific chocolate that Oscar has in mind. (A8)
> It's *the for specific candy. (A9)
> I'm not sure, but I think it's *the. (A1)
> *No article is used after *half of*. (A6)

In this example, the primary reason for the non-target-like article selection was the notion of specificity. Several advanced participants perceived the need for the definite article because the referent was a specific entity. However, they failed to recognize that a specific singular referent

can be accompanied by *a* when it is not known to the hearer/reader.

Finally, in non-referential indefinite contexts, seven advanced participants chose *the* instead of *a* in the following instance and provided a variety of reasons:

(66) Answer: (…) ate them all while he read This Week, (44) a popular weekly news magazine.
      Reasons: It's **the* for a well-known magazine. (A3)
                 It's **the* to describe the specific magazine. (A4, A6)
                 It's **the* for a specific magazine mentioned earlier. (A7)
                 It's **the* for a specific thing. (A10)
                 I think *a* and *popular* are not used together, so it's **the*.(A9)
                 I'm not sure, but I think it's **the*. (A5)

This particular item was also troublesome for the beginners and intermediates. In the case of the advanced group, specificity was a primary factor for the inappropriate article selection. As discussed earlier for item (71), specificity was often associated with the definite article by learners of English. However, when specificity was used as the only criteria for article choice without considering the hearer's knowledge, this association may have resulted in inaccurate article choices as in the above examples.

This section examined metalinguistic hypotheses that advanced participants employed in making article choices in the fill-in-the-article test. We find that advanced participants utilized intuitive judgments more frequently than beginners and intermediates as indicated by high frequencies of non-generalizable reasons for target-like article selections. The advanced group's performance in the fill-in-the-article test was not found to be at the level of the NES group, but we found evidence from the think-aloud data that much of the advanced participants' understanding of articles has become internalized.

(4) The NES Group

   In the previous sections, article choices and reasons for such choices were examined in details for the learner groups at three proficiency levels. Research Question 4 addressed the differences between the explicit knowledge of Korean learners of English as opposed to that of native English speakers. To respond to this question, I compared the metalinguistic hypotheses of English articles developed by Korean learners of English with those of native English speakers. For this purpose, a group of five native English speakers (NES) participated in the think-aloud protocol. The data from the think-aloud protocol are discussed in this section. The organization of this section is similar to those of the learner groups.

   First, distribution of specific reasons and non-specific reasons is presented in the following table.

Table 5. 15. Frequency of specific reasons and non-specific reasons in the NES group

	N	Percentage
Specific reasons	198	55.0%
Non-specific reasons	162	45.0%
Total	360	100.0%

   In this table, we note that the NES participants provided specific reasons (55.0%) relatively more frequently than non-specific reasons (45.0%). The percentage of non-specific reasons that the group produced is found to be considerably higher than the learner groups (21.6% in the beginner group, 12.1% in the intermediate group, and 27.4% in the advanced group). That is, the NES participants resorted to intuitive judgments rather than particular grammar rules. This observation seems to suggest that their knowledge of English articles is implicit. Anderson (2005, as cited in Roehr, 2007) explains that implicit knowledge is knowledge that cannot be brought into awareness or articulated. As such, the NES participants often failed to

provide explicit reasons for their article selections.

Next, accuracy for article choices based on specific reasons was compared with accuracy for articles selected with non-specific reasons.

Table 5. 16. Accuracy of article choices with specific reasons and non-specific reasons

	Accurate article choice	Inaccurate article choice	Accuracy
Specific reasons ($N = 198$)	190	8	96.0%
Non-specific reasons ($N = 162$)	150	12	92.6%

$p < .0001$

In this table, we note that the rates of accuracy for article choices in the NES group were relatively high and similar when they were made with specific reasons and when they were made without specific reasons. The articles selected with specific reasons were accurate in 96% of the time while those with non-specific reasons were accurate in 92.6% of the time. As discussed for the previous table, the high percentage of non-specific reasons was indicative of the implicit nature of the knowledge of articles assumed by NES participants. When intuitive judgments were applied, such article selections tended to be accurate.

Next, the reasons provided by the NES participants were examined in a similar manner with those presented by the learner groups. First, reasons that the NES group offered when choosing *the* accurately are presented in the following figure.

Figure 5. 10. Distribution of reasons provided by NES participants when *the* was selected appropriately
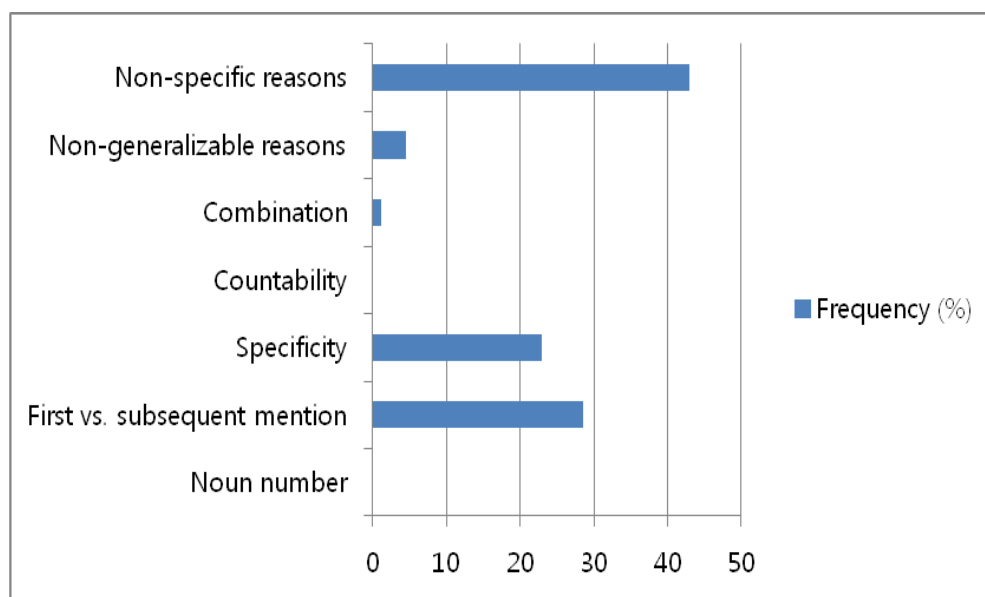
In this figure, we find that the most frequent reasons for the selection of the definite article were non-specific reasons. Specifically, the following expressions were frequently used to indicate the employment of intuitive judgments:[14]

(67) NES1: sounds better… just the way English is used
NES2: makes sense
NES3: sounds better… make sense
NES4: try to come up with hypotheses on the spot
NES5: I just know it is… I don't know why..

These phrases expressed by the NES participants suggest that they knew what articles to select but had difficulty in providing explicit explanations. As native speakers of English, the NES participants possess competence in the language, and this kind of competence is considered to be implicit in nature. In comparison, Ellis (2008) explains that metalingual knowledge must be learned through instruction or observation (as cited in Akakura, 2009). Without explicit

---

[14] The NES group used English in the think-aloud protocol.

instruction on English grammar, NES participants might not be equipped with all the technical and linguistic terminology necessary to explain their article choices explicitly.

When the NES participants provided specific reasons for the accurate selection of *the*, the most frequently cited factors were first vs. subsequent mention and specificity as in the following examples:

(68) Answer: When (26) the water in (27) the pot began to boil, (…)
     Reason: *The* is needed because it was mentioned before. (NES1)

(69) Answer: He added two cups of (37) pasta to (38) the boiling water and set (39) the timer.
     Reason: It's *the timer* because it's referring to a specific timer that he has. (NES4)

As in example (68), the first vs. subsequent mention distinction is applicable to anaphoric definite contexts. NES participants identified noun referents that were mentioned subsequently after the initial introduction and accurately supplied the definite article in (68). In addition, as shown in example (69), several NES participants based their choice of *the* on the feature of specificity.

The next figure illustrates frequencies of reasons provided by NES participants when they accurately selected the indefinite article.

Figure 5. 11. Distribution of reasons provided by NES participants when *a* was selected appropriately
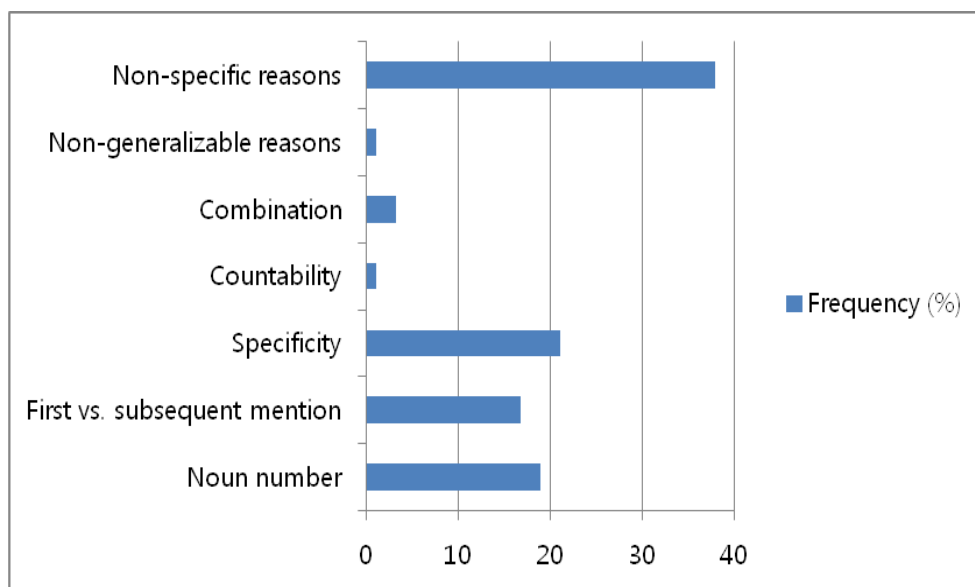
Similar to the case of the definite article, the NES participants provided non-specific reasons most frequently when they selected the indefinite article accurately. Again, this observation suggests that the knowledge of articles by native speakers is implicit.

In addition, the reasons related to specificity, noun number and first vs. subsequent-mention distinctions were cited with similar frequencies by the NES participants when they selected *a* appropriately. Examples of reasons are provided in the following:

(70)   Answer: Then he poured (69) cold nonfat milk into (70) a large glass.
         Reason: He chose a large glass and it's not a specific glass. (NES2)

(71)   Answer: (…) then used (24) a squeezer to squeeze all the lemon juice (…)
         Reason: It's *a squeezer* because he's just talking about one squeezer. (NES4)

(72)   Answer: (…) then put (50) the pasta into (51) a bowl.
         Reason: You need *a* here because it's new. (NES1)

In example (70), *a* was chosen because *glass* is not a specific one. In example (71), singularity of the referent was the reason cited for the selection of the article. In the last example, the

242

participant cited first-mention of a referent as a reason for the use of *a*.

Next, we examine reasons provided by NES participants for the target-like selection of $\emptyset$.

Figure 5. 12. Distribution of reasons provided by NES participants when $\emptyset$ was selected appropriately
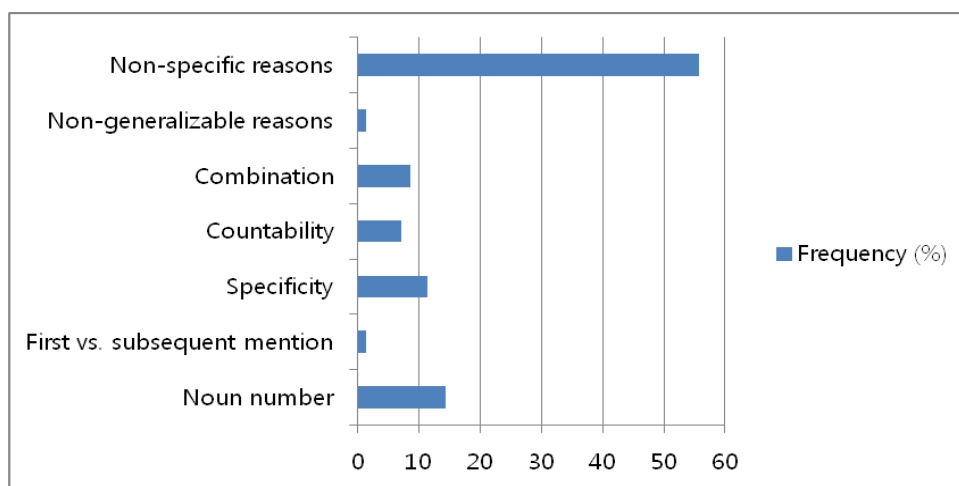
The figure shows that NES participants tended to select $\emptyset$ accurately but did not provide specific reasons. This observation suggests that their implicit knowledge was applied primarily in instances in which an article was not required. Other factors such as noun number, specificity, combination and countability were not as frequently expressed as found in the learner groups.

Up to this juncture, we have investigated the reasons behind target-like article selections by the NES group. The most prominent observation in this group was that NES participants relied mostly on intuitive and implicit judgments in making articles. For instance, the most frequent reasons for all three article choices were non-specific reasons, which refer to reasons based on intuition and feel. In addition to intuitive judgments, subsequent-mention and specificity were cited as factors that contributed to the selection of *the*. With regard to the indefinite article, non-specificity, singularity and first-mention were reasons for the accurate

choice of *a*. When NES participants selected ∅, they primarily provided non-specific reasons. The next frequent factors were plurality and non-specificity.

Now we turn to non-target-like article selections by NES participants. The NES group exhibited a nearly perfect performance and non-target-like article choices were rare. As such, only one instance was found to be problematic to NES participants, which is shown in the following table:

(73) Answer: Then he poured (55) the lemon juice onto (56) the pasta.
      Reasons: These nouns cannot be plural. When there is no use of unit, it takes
                *no article (NES1)
                It's *no article because no certain amount of lemon juice was
                mentioned (NES2)
                *No article sounds right. (NES3, NES5)

Four out of five NES participants did not seem to recall that *lemon juice* was already introduced into the story. Only one NES participant (NES4) remembered it and said, "This is the lemon juice that he had referred to way back with a squeezer". Instead, the reasons provided imply that they understand the uncountable noun status of *lemon juice* unless there is a unit of measurement (NES1) or a specification of its amount (NES2). Thus, they chose ∅. The other two participants (NES3, NES5) chose ∅ based on their implicit knowledge of articles. Without the previous mention, ∅ would have been the correct answer.

In brief, in this section, the think-aloud data obtained from the NES group were examined. The results point to a tendency of NES participants to select articles based on intuitive judgments.

5.2.4. Summary of key findings: think-aloud task.

In the previous sections, explanations that the participants provided for their article

choices in the fill-in-the-article test were examined. These verbally reported reasons allowed us to understand the kinds of explicit understanding of English articles that the participants have as opposed to the kind of understanding that the native English speakers possess. In this section, key observations of explicit knowledge of articles are summarized according to proficiency groups. First, the beginners exhibited the following tendencies with regard to reasons for explicit article choices.

(1) The beginner group provided specific reasons more frequently than non-specific reasons.

(2) Articles selected with specific reasons were more accurate than those selected with non-specific reasons.

(3) When beginners chose *the* accurately, the most frequently cited factor was subsequent-mention, followed by specificity and non-generalizable reasons.

(4) When beginners chose *a* accurately, the most frequent reason was singularity, followed by first-mention and non-specific reasons.

(5) When beginners chose ∅ accurately, countability was the most frequently cited factor, followed by non-specific reasons and plurality.

In turn, the following observations can be made about the kinds of reasons provided by intermediate participants for their article selections.

(1) The frequency of providing specific reasons was considerably higher than that of providing non-specific reasons.

(2) Articles chosen with specific reasons were more accurate than those selected with non-specific reasons.

(3) When intermediates selected *the* accurately, subsequent-mention was the most

frequently cited reason. Other minor reasons include non-specific and non-

generalizable reasons.

(4) When intermediates selected *a* accurately, the most frequently cited factor was

singularity, followed by first-mention and combination.

(5) When intermediates chose $\emptyset$ accurately, countability was cited most frequently,

followed by combination and non-specific reasons.

Next, patterns of explicit reasons reported by the advanced participants for their article

selections can be summarized as the following:

(1) Specific reasons were provided more frequently than non-specific reasons.

(2) Rates of accuracy with which articles were selected with specific reasons were

similar to those with which articles were selected without specific reasons.

(3) When advanced participants selected *the* accurately, the most frequent reason was

subsequent-mention, followed by non-specific reasons and non-generalizable reasons.

(4) The majority of instances in which the advanced participants selected *a* accurately

were based on singularity of nominal referents.

(5) When advanced participants selected $\emptyset$ accurately, the most frequent reasons were

non-specific. Other minor reasons include plurality.

Finally, the NES group can be characterized as the following with regard to their article

choices and explanations for such choices.

(1) The frequency of specific reasons was slightly higher than that of non-specific

reasons.

(2) Rates of accuracy were similar between articles selected with specific reasons and

those selected without specific reasons.

(3) The majority of accurate selections of *the* were accompanied with non-specific

reasons. Other minor factors included subsequent-mention and specificity.

(4) The most frequent reason for the accurate selection of *a* was non-specific reasons.

Other reasons include specificity, singularity and first-mention.

(5) The majority of instances in which ∅ was accurately selected were based on non-

specific reasons.

In this previous section, explicit knowledge of articles was examined by looking at the think-aloud data obtained from the participants. In the next section, explicit knowledge of articles is compared with implicit knowledge in three case studies to explore interactions between the two kinds of knowledge of the English article system among Korean learners of English.

5.3. Comparisons between Explicit Hypotheses and Implicit Use of English Articles

5.3.1. Introduction.

In this study, both implicit and explicit knowledge of English articles were investigated for Korean learners of English at three proficiency levels. Specifically, data on the implicit knowledge of English articles held by Korean learners of English were gathered through oral and written narratives and analyzed quantitatively with regard to English language proficiency and genres of language production. The results of the analyses on the implicit use of English articles were also reported in Chapter IV. Data obtained from the explicit knowledge of English articles that the participants possess were obtained from the grammaticality judgment task and the think-aloud task. Results of the quantitative and qualitative analyses of the data on metalinguistic knowledge were also discussed in this chapter.

In the remaining part of this chapter, patterns of spontaneous use of English articles of three selected participants in their oral narratives are discussed in comparison with the explicit hypotheses they provided in the think-aloud protocol. Comparisons between the implicit and explicit knowledge of English articles were made to address the final research question that sought to determine how learners' metalinguistic knowledge of articles is operationalized in their oral production.

The three participants selected for this part of the chapter are one beginner, one intermediate, and one advanced speaker. In selecting candidates for this comparative analysis, participants who were situated in the middle of respective proficiency groups were considered. The scores and accuracy rates of all the participants are available in Appendix F. The objective of this section is to provide qualitative illustrations of how implicit and explicit knowledge may diverge or converge for different types of articles and NP contexts in a case-study format. I consider that the three participants whose language was extracted from the larger corpus may not be representative of each of their groups since the amount and kinds of language input, the length of stay in the U.S. and other non-linguistic variables may mediate the use of the linguistic variable under investigation. Thus, the three case studies are reported here as examples of how the understanding and use of articles may be operationalized in the explicit and implicit language tasks.

The organization of this section is as follows. First, metalinguistic hypotheses about English articles provided by the participants are compared with the English articles used in their oral narratives. These comparisons are reported by individual participants. Then, findings from the three cases are discussed to gain insights about the differences and similarities between explicit and implicit knowledge of English articles

5.3.2. The Beginner Speaker (B4): a qualitative examination of article use.

At the time of the data collection, the selected beginner speaker was staying with a Korean family as a short-term missionary. The total length of stay in the U.S. was three months and the age of arrival was 23 years old. According to the sociolinguistic interview, the participant learned English in school for six years as part of regular school curriculum, which mainly focused on reading comprehension and grammar. At the time of the study, she was attending a full-time ESL program offered by a state university, but she revealed in the interview that she did not have much opportunity to use English outside of the classroom since she was staying with a Korean family.

In the fill-in-the-article test, the beginner speaker scored 81.1% accuracy in instances in which *the* was required and 85.0% accuracy in instances in which *a* was required. Metalinguistic hypotheses on the use of English articles expressed by the beginner speaker in the think-aloud protocol can be summarized as the following.

(1) The definite article (*the*) is used when the referent has been previously mentioned and/or when the referent is specific.
(2) The indefinite article (*a*) is used when the referent is introduced for the first time, when the referent is countable, and/or when the referent is singular.
(3) No article is used when the referent is uncountable or plural.

These hypotheses suggest that the rule concerning first-mention and subsequent-mention is an important part of the participant's explicit knowledge of article. In other words, the learner selected *the* when the referent was identified as having been previously mentioned and choose *a* when the referent was not mentioned before. As White (2010) reports, the choice of articles based on first and subsequent mentions of referents is introduced in most ESL textbooks as a primary rule for article use and this factor seems to have conditioned her use of that rule.

249

Another hypothesis that the beginner frequently applied in article selections is related to countability and noun number. That is, *a* is the article of choice for a countable, singular referent and no article is used for an uncountable or plural noun. These rules of article use based on countability and noun number are also presented in ESL textbooks and taught in ESL classes.

The think-aloud data obtained from the beginner speaker and the hypotheses she generated were mainly based on simple rules and were applied without sensitivity to relevant discourse contexts. These rules may be applicable to anaphoric definite contexts, but in associative definite contexts, several discourse factors need to be considered. However, the beginner failed to consider contextual factors, such as the fact that the setting of the story is a kitchen, resulting in article errors. Some examples of how the beginner applied these rules without considering contextual features are shown in the following:

(74) Answer: Oscar took (41) a cucumber out of (42) the fridge.
     Reason: I think you need *a* here because it is a new noun.

(75) Answer: He added two cups of (37) pasta to (38) the boiling water and set (39) the timer.
     Reason: I choose *a* because it's mentioned for the first time.

In the above examples, the beginner learner chose *a* based on the judgment that these noun referents were not mentioned before. However, in the context of Oscar's kitchen where this story took place, certain kitchen tools and appliances such as *fridge* and *timer* are typically considered uniquely identifiable when the speaker and the hearer share cultural understating, thus taking the definite article. The beginner speaker did not consider this contextual aspect and mainly depended on structural cues to determine whether or not noun referents were previously mentioned in a technical and strict manner. Butler (2002) also reported that learners at lower proficiency levels tend to develop context-insensitive hypotheses and that these learners tend to

250

"rely heavily on structural cues without considering more dynamic contexts" (p. 466). In the case of the beginner participant, the discourse context of the kitchen was not considered in determining whether or not *fridge* and *timer* should be considered definite. The unconditional application of the rules concerning first-mention and subsequent-mention resulted in the incorrect selection of *a*.

Another concern regarding how the beginner applied hypotheses about English articles is related to noun countability. In most instances, the participant did not explicitly cite the singularity of a noun referent as a reason for the choice of *a*. Instead, first-mention was most frequently cited as a reason for the selection of the indefinite article. Still, the selection of *a* implies that the speaker considered countability and singularity of a noun referent because the think-aloud data revealed that the participant was indeed aware that *a* cannot be used for plural or uncountable nouns. Non-target-like choices of *a* were made when the participant failed to perceive the countability of noun referents accurately, as in the following examples:

(76) Answer: He added (52) a heaping tablespoon of (53) butter to (54) the hot pasta and let it melt in.
        Reason: I think you need *a* because it's a new noun.

(77) Answer: He sprinkled (62) garlic powder and grated (63) parmesan cheese onto (64) the pasta.
        Reason: You need *a* because you can count cheese.

In example (76), the participant cited new information, the information structure of a clause, as a reason for the choice of *a*, but the fact that *a* was chosen instead of no article implies that the participant perceived *butter* as a countable noun, even though it was not explicitly stated. In example (77), the beginner explicitly stated that *a* was selected because *parmesan cheese* was a countable noun. Similarly, this participant also determined that nouns such as *pasta*, *small print*, *lemon juice* and *gas* were countable. These examples suggest that the beginner have difficulties

251

in determining countability of noun referents based on respective contexts.

In short, the beginner participant has a basic understanding of how articles tend to be in English, but her metalinguistic knowledge is limited in the sense that discourse and context related factors are not considered properly in the use of articles. Further, the perception of definite status of referents in associative definite contexts was challenging to the participant. Pica (1983) argued that this challenge may be attributed to lack of exposure to authentic use of articles, which is essential for L2 learners to develop appropriate conceptualizations of articles.

Turning to the participant's implicit knowledge, the oral narrative produced by the beginner was short in length and did not include many instances in which articles were produced. Instead, possessives were used frequently throughout the narrative. As Butler (1999) pointed out, spontaneous production data may pose a challenge to researchers in SLA because of the possibility of avoiding the use of difficult grammatical features. In this respect, the beginner tended to use possessives in avoidance of using articles. In instances in which articles were required, *the* was produced more accurately (69.2%) than *a* (30.4%) in the spontaneous production.

In line with the metalinguistic hypotheses discussed above, the beginner used the definite article relatively accurately when noun referents were mentioned subsequently, as in the following:

(78) There was a big retreat for mission. I participated in the retreat and I vowed to God.

In the oral narrative, the number of anaphoric references was small, but it was clear that the beginner was able to use *the* in subsequent-mention contexts. As discussed in the earlier section, the first-mention and subsequent-mention operates as a guiding rule since ESL textbooks

resort to this rule.

Nevertheless, omission of articles was frequent in the beginner's oral narrative. For instance, nominals in the following examples were already introduced in the context and thus required the definite article:

(79) God call me through *pastor, but I want more.. more exactly (…)
(80) I didn't have power in *family, so I obeyed their words (…)
(81) I dreamed.. in *dream, my sister praying.. my sister is praying (…)

In the above examples, *the* was omitted in these anaphoric definite contexts in the beginner's oral narrative. In the fill-in-the-article test, omission was not an option because it was a multiple-choice test that directed participants' conscious attention to article use. In comparison, we find that the beginner tended to omit the definite article, which was possibly mediated by cognitive constraints related to spontaneous production.

A similar tendency to omit *a* was also attested in the beginner's oral narrative. In the following examples, *a* was omitted for NPs such as *short-term missionary*, *mission*, and *job* as they were used in indefinite contexts:

(82) The story is about *short-term missionary.. how can I decide to come as *short-term missionary
(83) In 2005, there is.. there was a big retreat for *mission.
(84) (…) they already got *job and so.. they want to me got *job or work.

In the above examples, the indefinite article *a* was not produced as required. It seems unclear whether the participant mistakenly omitted *a* under pressure of spontaneous speech or whether the countability of these noun referents was judged inaccurately. According to the metalinguistic hypotheses, the participant selected *a* based on such features as first-mention and singularity of noun referents. Thus, the omission of *a* in the above examples may suggest that the beginner perceived nouns such as *short-term missionary*, *mission*, and *job* as uncountable nouns. This

possibility is further supported by the fact that *short-term missionary* was produced without the required indefinite article *a* in all seven instances in the short narrative.

Even though the beginner omitted *a* in several instances, *a* was produced accurately in certain structural constructions. For instance, *a* was used accurately when a noun referent was in the predicate position of the copular *be* verb, as in the following examples:

(85)   I <u>am *a* member</u> in JDM.
(86)   When I <u>was *a* freshman</u> in university, (…)
(87)   There <u>was *a* big retreat</u> for mission.
(88)   When I <u>was *a* senior</u>, I worried about my future.

In these examples, we note the target-like use of *a* together with the verb *be*. The use of *be* together with an adjectival predicate or a nominal predicate is commonly introduced to English language learners at an early stage of instruction. As such, the beginner may have been familiar with this type of construction and able to use *a* in a way similar to the use of formulaic expressions.

While certain tendencies were evident in the oral production of articles, fluctuations were also observed in the ways the beginner sometimes produced articles and omitted them in other instances. Let us examine the following examples.

(89)   When I had a quiet time, <u>*bible</u> say me, if you have.. if you do.. if you vow to God, you.. you do that. When I saw <u>the bible</u>, I afraid of God, so I decide again.. I will go, some day (…)

(90)   When I had <u>a quiet time</u>, bible say me.. (…) but God say me again quiet time.. in <u>*quiet time</u>.. so I again decide (…)

In example (89), we note that the participant omitted *the* in the first instance, but produced the article in the second instance. Similarly, in example (90), *a* was produced in the first instance, but later the article was omitted in the second instance. In second language

254

acquisition, fluctuations are commonly observed phenomena and they provide evidence for constant development and restructuring of learners' inter-language.

In summary, the metalinguistic data suggest that the beginner has a basic and incomplete understanding of articles as evidenced in the absence of considerations for contextual factors in making article choices. In the fill-in-the-article test, the participant was able to select *the* for subsequent-mention of referents and *a* for first-mention of singular referents. In other instances, the majority of non-target-like article choices were due to inaccurate perception of noun countability and inability to recognize discourse factors in the determination of definiteness. In the oral narrative, articles were produced accurately in several instances. Also, fluctuations between target-like and non-target-like article production were observed, which suggest that the beginner's implicit knowledge is in developmental stages. However, article omissions were frequent for both articles in a variety of contexts, an outcome which may be attributed to cognitive and linguistic constraints that L2 speakers at lower levels of proficiency experience in spontaneous oral production.

5.3.3. The Intermediate Speaker (I5): a qualitative examination of article use.

The intermediate participant selected for this section had lived in the U.S. for a little more than two years at the time of the study. He came to the U.S. after finishing high school to enter an American college. When data were collected, he was studying business and accounting at a community college and was planning to transfer to a university.

The intermediate speaker outperformed the beginner in the fill-in-the-article test. The definite article *the* was selected accurately in 81.1% of the instances, and the indefinite article *a*, 95.0%. Metalinguistic hypotheses provided by the intermediate participant can be summarized as

the following:

(1) The definite article (*the*) is used when the referent has been previously mentioned and/or when the referent is specific.
(2) The indefinite article (*a*) is used when the referent is singular, unspecified and singular, or is not previously mentioned and singular.
(3) No article is used when the referent is not previously mentioned and plural, or is not previously mentioned and uncountable.

With regard to the use of *the* and *a*, the intermediate speakers selected the articles based on whether or not a given referent was previously mentioned or not. Further, the intermediate speaker seems to hypothesize that the feature of specificity is an important factor in article usage. In other words, the metalinguistic hypotheses developed by the intermediate speaker suggest that *the* is used for a specific and particular referent while *a* is used for an unspecified referent. Another important feature for the use of the indefinite article is whether or not the referent is singular, which also implies that the given referent is countable. If it is a singular referent, *a* is used. The intermediate participant selected Ø when the referent was plural or uncountable and was not previously mentioned.

The participant faced little challenge in detecting subsequent-mentions and selecting *the* accordingly. In comparison, associative definite contexts posed more of a challenge. Similar to the beginner in the previous segment, the intermediate speaker also had difficulties in properly recognizing associative definite contexts. The following examples are illustrative of such difficulties:

(91) Answer: On (34) the side, in (35) small print at (36) the very bottom, was "Product of U.S.A."
    Reason: I choose *a* here because *bottom* was not mentioned before and it is singular.

(92) Answer: While (40) the pasta was cooking, Oscar took (41) a cucumber out of (42) the fridge.
    Reason: Because *fridge* is mentioned for the first time and it is singular, it

should be *a*.

In example (91), the definiteness of *bottom* is determined in association with *a box of pasta* that was mentioned before. When a box is introduced, certain attributes of the box are considered definite and can be uniquely identified by the speaker and the hearer. Therefore, in such associative definite instances, the definite article is required. Similarly, the setting of the entire story is Oscar's kitchen, and by way of association with the kitchen, certain appliances and items commonly observed in a typical kitchen are considered definite. Therefore, *fridge* in example (92) takes the definite article. However, the intermediate speaker failed to examine the larger context to determine the definiteness status of a referent based on the contextual factors.

The choice of *a* was produced accurately in the fill-in-the-article test. By applying hypotheses based on singularity and first-mention, the intermediate speaker successfully selected *a* in 95.0% of the instances in which *a* was required. However, the participant exhibited a small but noticeable tendency to over-use *the* based on the feature of specificity, as in the following examples:

(93)   Answer: These studies are based on (3) experiments with (4) mice and other
                 small animals.
       Reason: I think you need *the* here because *experiments* refer to *these*
                 studies and thus they are specific.

(94)   Answer: (…) ate them all while he read This Week, (44) a popular weekly
                 news magazine.
       Reason: It's *the* because it is a specific magazine.

In these examples, specificity played a role in the selection of *the*. L2 learners' tendency to associate specificity of a referent with the definite status has been previously documented in the literature (e.g., Hawkins et al., 2006; Ionin et al., 2004; Liu and Gleason, 2002, Zegarac, 2004). Indeed, as presented earlier, the hypotheses about the definite article reveal that the specificity of

a referent is considered an important criteria for the selection of *the*. As such, when a referent is specific but is not known to the hearer, the intermediate participants tends to choose *the* inaccurately.

Up to this juncture, a metalinguistic understanding of the intermediate speaker was examined based on explicit hypotheses and reasons provided by the participant in the think-aloud protocol. These hypotheses are compared with patterns of article use exhibited in the oral narrative.

There were only a few instances in which *the* was required because the participant mainly talked about future plans and goals, which tend to yield more indefinite instances. Due to the small number of instances in which *the* was required and a few article errors, the rate of accuracy in terms of percentage appeared relatively low.

The oral narrative yielded a greater number of indefinite contexts. The intermediate speaker produced *a* only in a half of the instances in which *a* was required. Below are examples of instances in which *a* was accurately produced:

(95) My dream is to be <u>a sports agent</u>.
(96) And I have another dream. I hope to be <u>a volunteer</u>.
(97) If they have <u>a very famous player</u>, they can get 10% of their income.
(98) (…) because there is <u>a good place</u> to learn something.

These excerpts exemplify how the intermediate participant can use the indefinite article accurately. In particular, *a sports agent* was produced accurately five times and *a volunteer*, six times, in this oral narrative.

However, in a number of instances, the tendency to omit the indefinite article was observed, as in the following examples:

(99) (…) but they already said, that is not <u>*good job</u>.
(100) They already say that I have to go <u>*big company</u> to earn money.

(101)   When I was *high school student, I already said, I wanna be a volunteer (…)
(102)   (…) they can get 10% of their income, so that is *good deal.

In the above examples, the indefinite article was absent. In particular, *big company* was used in all four instances without *a*. Unlike omissions in the oral narrative, the participant exhibited a firm understanding of the indefinite article in the think-aloud protocol and selected *a* in the fill-in-the-article test with very high accuracy. Therefore, frequent omission of *a* in the oral narrative may be in part explained in terms of constraints of spontaneous production. Phonologically, the indefinite article is a one-syllable word that is unstressed. Semantically, this article does not seem to contribute much to the overall interpretation of the message. Hence, in a spontaneous speech, the learner may fail to express it. In this regard, it is possible to posit that conceptual changes concerning the construction and interpretation of nominals in English have not been completed by this intermediate participant. Without conceptual restructuring for target-like nominal expressions, a learner may easily omit articles under pressure to produce L2.

In addition to omissions, fluctuations were also observed in the use of the indefinite article, like the following examples:

(103)   my parents already said, if you wanna be a volunteer, you have to earn a lot of money before you can be *volunteer.
(104)   actually, I hope to establish a big soccer school (…)
(105)   I also establish *soccer school, like *sports school at Jeju Island in Korea.

In example (103), fluctuations were observed as attested in the use of *a* which was used in the first instance but omitted in the second instance even though it is very similar sentence structure. Similarly, in example (104), *a* was accurately produced but the article was absent in example (105). These tendencies to use articles in some instances and omit them in other instances are indicative of the on-going restructuring and conceptual changes in the learner's grammar of English.

Another interesting characteristic with regard to article use in the oral narrative is the occurrences of *some* in place of articles in indefinite contexts, as in the following examples:

(106)   I wanna give *<u>some chance</u> to learn sports to poor people.
(107)   My father is working.. working.. managing *<u>some factory</u>, so (…)
(108)   So I'm looking for *<u>some job</u>.. maybe I hope to work big company first (…)

In these examples, *some* was used instead of *a*. In fact, several scholars include *some* as a member of the article system. For instance, in proposing a unified conceptual framework of articles, White (2010) treated unstressed *some* as an article. The use of *some* can also be explained as an avoidance strategy. That is, as proficiency increases, an L2 speaker may develop more linguistic devices and strategies and thus tend to avoid using certain linguistic features that they have difficulties with. In this sense, the use of *some* in the above examples may be a strategy that the intermediate speaker employed in spontaneous oral production.

In short, the think-aloud data revealed that the intermediate participant was fluent in meta-language to explicitly describe reasons for article choices that were made with high accuracy in the fill-in-the-article test. In comparison, omissions and fluctuations were observed in the use of articles in the oral narrative. Lower accuracy with which articles were employed in the oral narrative may point to the effect of production-related constraints as well as an on-going development of implicit knowledge necessary for proficient use of articles in spontaneous production.

5.3.4. The Advanced Speaker (A6): a qualitative examination of article use.

The advanced speaker selected for the comparison between implicit and explicit knowledge of articles was attending a university at the time of data collection. He immigrated to the U.S. with his parents at the age of 15. He had been living in the U.S. for seven years when

this study recruited him as a participant.

In the fill-in-the-article test, the advanced speaker scored 83.8% in instances in which *the* was required and 90.0% in instances in which *a* was required. The metalinguistic hypotheses that emerged from the think-aloud data can be summarized in the following manner:

(1) The definite article (*the*) is used when the referent has been previously mentioned.
(2) The indefinite article (*a*) is used when the referent is singular.
(3) No article is used when the referent is plural.

These rules seem to be basic, but they represent only those reasons explicitly stated by the participant. With regard to the definite article, as the hypothesis indicates, the participant was able to identify previously mentioned referents and selected *the* accurately. In case of associative definite contexts, fluctuations were observed. In some instances, the participant used intuitive judgments and selected *the* accurately while providing non-generalizable reasons such as "sounds better" and "sounds more natural", as in the following examples:

(109)   Answer: He filled (13) a pot with (14) water, put (15) the pot on (16) the stove, and turned on (17) the gas.
        Reason: I think it is *the*. It sounds better.

(110)   Answer: On (34) the side, in (35) small print at (36) the very bottom, was "Product of U.S.A."
        Reason: In both places, *the* sounds right.

Comments as the ones cited above suggest that much of the participant's knowledge of English articles is already internalized as implicit knowledge, which cannot be easily explained. This observation may be related to the characteristics of language learning that the participant has experienced. Like most advanced participants in this study, this participant went to high school in the U.S. and most of his L2 learning has been more implicit than explicit. In the sociolinguistic interview, he said that he attended ESL classes in high school but much of his English language learning experiences have been mediated by exposure to natural language use and social

interactions with native speakers. Thus, the participant's knowledge of English articles may be implicit and difficult to verbally express.

With regard to indefinite contexts, the participant seems to have developed a hypothesis that *a* is used for a singular referent and no article is used for a plural referent. In most instances, this distinction worked successfully. However, in other cases, the participant was found to be affected by this hypothesis. As a result, there was a tendency to select *a* or $\emptyset$ inaccurately in associative definite contexts simply based on noun numbers of referents. The following examples are illustrative of this tendency:

(111)　Answer:　(…) he grabbed (28) a box of (29) pasta from (30) the cabinet.
　　　　Quote:　I think it is *a because *cabinet* is singular.

(112)　Answer:　He peeled it, sliced it, salted (43) the slices, and ate them all (…)
　　　　Quote:　*No article is needed here because *slices* is plural.

(113)　Answer:　When Oscar finished (67) the pasta, he washed (68) the dirty dishes and utensils.
　　　　Quote:　*Dishes* is plural, so I will *not put any article here.

In example (111), *a* was used because of the perception that *cabinet* is a singular nominal. Similarly, the advanced participant chose *a* for *stove*, *timer*, and *fridge*. This observation is indicative of how the advanced participant may not have been able to accurately recognize the role of contextual factors on article use in English.

Now, we turn to examine how the advanced participant used articles in his oral narrative. In speaking, *the* was produced with an accuracy of 94.7%, and *a* with an accuracy of 76.9%. The fact that the accuracy in the use of *the* was higher in the oral narrative than the fill-in-the-article test suggest that the participant has a firm understanding regarding implicit knowledge of the definite article. As such, the oral production of *the* was fluent and accurate.

262

While *the* was produced accurately in contexts that required the article, the tendency to over-use *the* in indefinite contexts was also attested in his oral narrative. This tendency may explain the lower accuracy for *a* (76.9%) in the oral narrative than in the fill-in-the-article test (90%). In the following examples, *the* was over-used in indefinite contexts:

(114)   That was the first time I have ever flied *the airplane.
(115)   That was the first experience studying in *the foreign middle school.
(116)   'I will give it a try', and I enrolled in *the flight school in Princeton.
(117)   I couldn't be a pilot in Korea because they require *the good eye vision.
(118)   Of course, I was in *the ESL classes because I was not really fluent in English.
(119)   Since I was in *the middle school, I had to come back to Korea for *the high school.

The above examples illustrate how the advanced participant extended the use of *the* in indefinite contexts. In the first three examples, *the* was produced instead of *a*, and in the last three examples, *the* was produced when an article was not required.

The comparison of this participant's hypotheses about articles and how articles were actually produced in the oral narrative revealed that the advanced participant has a stable knowledge of English articles. Articles were chosen accurately in the fill-in-the-article test, while articles were used fairly accurately in the oral production. While the participant did not have much difficulty in using articles in a natural speech, he expressed during the think-aloud protocol that he did not have much confidence in the use of articles. This may be related to the fact that his English language acquisition was done mostly in an implicit manner and the participant felt challenged in trying to provide explicit reasons behind article choices. This characteristic of the advanced participant is in contrast with the intermediate participant who was confident in explicitly explaining reasons for article choices but had lower accuracy in the actual use of articles in the oral narrative.

Up to this juncture, metalinguistic hypotheses provided by the three participants were

compared with how they employed articles in their oral narratives. The comparisons were conducted to understand the differences between explicit knowledge revealed in explicit hypotheses and implicit knowledge manifested in oral production of articles.

In the next chapter, results of the current study are summarized and discussed in view of the research questions this study was set to address. Also, theoretical and pedagogical implications of the current study's findings are presented.

Chapter VI. Discussion

6.1. Introduction

The goal of the current study was to present a comprehensive view of how Korean learners of English understand and use the article system. To investigate the knowledge and use of articles by Korean learners of English, five research questions (RQs) were formulated in the study. RQs 1 and 2 focused on how articles were used by the participants in the oral and written discourses. RQs 3 and 4 were concerned with the explicit knowledge of articles exhibited by the participants in the think-aloud exercise in comparison with that of native English speakers. RQ5 addressed how the implicit knowledge of articles reflected in oral and written article production was compared to the explicit knowledge of articles evidenced in the think-aloud data.

To address the five research questions, four tasks were performed by 30 Korean learners of English, who were divided into three proficiency groups. The tasks employed in this study included the oral narrative task, the written narrative task, the grammaticality judgment task, and the think-aloud task. In addition, five native English speakers undertook the grammaticality judgment task and the think-aloud task to serve as a comparative basis. The quantitative and qualitative data gathered from these tasks were analyzed to explore the implicit and explicit

knowledge of English articles that the participants possess. The results of the analyses were reported in detail in Chapters IV and V.

In this chapter, findings of the current study are discussed as they relate to the research questions. The organization of this chapter is as follows. Firstly, key findings with regard to article use in oral and written narratives are discussed as they respond to RQ 1 and RQ 2, which aim to investigate patterns of article production in the oral and written discourses. Secondly, learner hypotheses of English articles obtained from the think-aloud protocol are discussed in light of the literature. This segment serves to address RQ 3 and RQ 4, which are concerned with the kinds of metalinguistic awareness of articles exhibited by Korean learners of English as well as native English speakers. Thirdly, the explicit knowledge expressed in the think-aloud protocol is discussed in comparison to the implicit knowledge of articles manifested in the oral narratives. This segment is intended to answer RQ 5 that focuses on how explicit knowledge of articles is reflected in the spontaneous oral production of articles. Finally, the chapter concludes by presenting theoretical and pedagogical implications of the current study.

6.2. Patterns of Article Use in Oral and Written Discourses

For the purpose of addressing the first and second research questions which focused on patterns of article use according to English language proficiency, the analyses of English articles produced in the corpus of the oral and written narratives were conducted. An examination of articles employed by the participants revealed their implicit knowledge of English articles at different proficiency levels. In this section, distinctive characteristics of article use that emerged in the analyses are considered. Patterns of article use in the oral narratives are discussed first, which is followed by a discussion of the written production of articles in comparison to their oral

production.

With regard to the production of *the* and *a* in the oral narratives, a positive correlation was found between the accuracy in which they were produced and the external variable of English language proficiency level. A positive correlation between the accuracy of article use and proficiency was previously documented by Huebner (1983), Oller & Redding (1971), Parrish (1987) and Young (1996), among others. More specifically, Huebner (1983) and Parrish (1987) adopted a longitudinal approach and documented patterns of article use by L2 learners over an extended period of time. Both studies reported that their participants began to use articles more accurately as their proficiency improved. Oller & Redding (1971) also reported a positive correlation based on the results from a multiple-choice test on article use. In this sense, the current study compliments the previous findings in that it has attested to an increased level of accuracy in the oral production of *the* and *a* in relation to an increase in proficiency for all participants.

Between *the* and *a*, the participants at all proficiency levels exhibited a tendency to produce *the* more accurately than *a* in their oral narratives. In fact, the accuracy for *the* was nearly twice as high as the accuracy for *a* in all groups. With regard to the order of acquisition of the two articles, several studies (e.g., Huebner, 1983; Master, 1987; Parrish, 1987; Thomas, 1989) have suggested early integration of *the* into the learner's interlanguage before *a*. The findings of this study indicate that participants at all proficiency levels are able to produce *the* more accurately than *a* with respect to spontaneous oral production.

Several explanations can be suggested for the low accuracy in the oral production of *a*. One possible explanation may be related to the complexity of the article system. The accurate use of *a* in English requires a consideration of a multitude of factors that are semantic, pragmatic,

266

contextual, and discursive. For instance, in the binary schema of English articles, Master (1990) suggested that speakers of English need to make three distinctions in order to produce *a* accurately: definiteness, countability, and noun number of NPs. Simultaneous determination of the three distinctions is a complicated task because each of these distinctions is not inherent in lexical items but highly dependent on a given discourse context. In other words, the same lexical item (e.g., *chicken*) can be definite when the referent denoted by the lexicon is already known to and identifiable by the hearer in the immediate discourse space, as in *Look at the chicken over there*. When the referent is assumed unknown to the hearer, the speaker needs to use the indefinite article, as in *I saw a chicken in my backyard yesterday*. Likewise, countability and number are also context-dependent distinctions. When the noun refers to chicken meat, it is considered uncountable, as in *I want to have chicken for dinner*. When it denotes an individual bird, it is considered countable, as in *I saw many chickens on the farm*. As suggested by these examples, there is not a simple form-to-meaning mapping for the indefinite article, and the accurate use of the form is dependent on a given context. As such, Korean learners of English may experience difficulty in accurately producing *a* in a spontaneous speech.

In addition, the marked difference in accuracy between the two articles may be related to the manner in which definiteness and indefiniteness are marked in Korean, the L1 of the participants in this study. According to Kim (1990), definiteness can be indicated in Korean by demonstratives such as *–geu* in some cases, but there is no morpheme equivalent of *a* that indicates indefiniteness in a similar way as English does. That is, Korean learners of English may associate *the* with the use of certain demonstratives in Korean, but they do not have a form in their L1 to which they can map the use of *a*. The different manners in which English and Korean express the notions of (in)definiteness may have thus conditioned the participants in this study to

produce *a* with lower rates of accuracy than *the*.

Difficulties associated to the accurate spoken production of *a* may have been further compounded by cognitive processing constraints inherent in spontaneous L2 production. In a spoken discourse, pressure to produce L2 in a spontaneous manner tends to constrain cognitive resources of the learner and lead to performance-related errors. Thus, as Trenkic (2009) maintained, determining all aspects of NPs in a particular discourse context for the accurate production of *a* in L2 oral speech may have imposed increased cognitive challenges to L2 speakers in this study.

In the oral narratives, the accurate use of *the* and *a* was affected by another linguistic variable, which is the presence or absence of adjectives in NPs. Participants at all proficiency levels produced *the* and *a* more accurately in instances in which adjectives were absent in NPs, a finding previously attested in Trenkic (2007). In an attempt to explain the reasons why adjectives seem to condition accuracy of article use, Trenkic proposed that L2 learners of English mis-analyze articles as nominal modifiers and treat them as lexical items. However, it is not clear whether the low accuracy rates of articles in [+ Adjective in NP] contexts in this current study are attributed to the mis-analysis of articles as nominal modifiers as argued by Trenkic (2007). Other possible factors such as cognitive processing concerns could condition these tendencies.

In this regard, a possible hypothesis to consider is that participants tend to omit articles more frequently when a pre-noun slot in an NP is already filled with an adjective. When adjectives are used to modify nouns, speakers need to pay attention to formulate an NP accurately, leaving less cognitive resources available for the accurate production of articles. In fact, Trenkic (2009) has also recently adopted information processing theory to position her findings within the ideas of limited and finite capacity for information processing in humans.

Thus, Trenkic (2009) explained processing constraints as the cause of article omissions in NPs that contain adjectives. In other words, when two forms (i.e., an article and an adjective) compete for the same pre-nominal slot, under processing constraints, L2 learners whose L1s do not have articles may choose to express the adjective and omit the article because the more semantically and phonologically salient form is likely to be chosen.

With respect to non-target-like production of articles in the oral narratives, article omissions were found to be frequent in all groups. Specifically, when *the* was not produced accurately, the form was mostly omitted. The tendency to omit an article may be indicative of an influence of the L1 on the article use by Korean learners of English. In this regard, several scholars have documented the effect of L1s on English article use when learners' L1s do not have an article system, and their studies point to mainly article omission (e.g., Liu and Gleason, 2002; Mizuno, 1999; Park, 2006; Robertson, 2000; Thomas, 1989; Young, 1996; Zegarac, 2004, among others). While article omission may persist in L2 use of articles, the current study revealed that the frequency of article omission decreased along the proficiency continuum. This finding suggests that L1 effect on article use is more evident at lower levels of proficiency and the effect tends to diminish as the learner's understanding of articles approximates the target language.

In the case of the indefinite article, different patterns of non-target-like use emerged depending on NP contexts: non-referential indefinite contexts (i.e., first mention of a specific referent unknown to the hearer) and referential indefinite contexts (i.e., reference to an unspecified entity). Specifically, *a* was omitted more frequently in non-referential indefinite contexts than in referential indefinite contexts. A possible explanation for the tendency to omit *a* for non-referential uses may be related to the virtual nature of indefinite reference in these

269

contexts. In this regard, Langacker (2008b) explains that the actual/virtual distinction offers a viable characterization of definite and indefinite nominals. That is, the default case for definites is that their referent is an actual instance of its type, while the default expectation for indefinites is virtuality of a referent. For example, in *My daughter played with the puppy after school*, the underlined NP denotes an actual pet. In contrast, in *My daughter wants a puppy for her birthday*, the underlined indefinite NP does not refer to an actual dog in particular but a virtual one. The virtual reference is a characteristic of non-referential indefinite contexts. In these contexts, learners of English may tend to omit *a* since they do not have an actual entity in mind to refer to, a finding evident in the frequent omission of *a* in the oral narratives in this study.

Instances in which an indefinite NP refers to a specific referent were categorized as referential indefinite contexts in this current study. In these instances, the reference denoted by the indefinite NP is perceived as actual in the mind of the speaker at the time of the utterance but still virtual in the mind of the hearer. For instance, in *I downloaded a fun on-line game yesterday*, the underlined NP refers to a specific and existential entity in the mind of the speaker. The actual and specific nature of a nominal referent in referential indefinite contexts may motivate an L2 English learner to favor the use of *the* instead of the required *a*. Indeed, the tendency for L2 learners of English to over-use *the* for specific referents has been reported by several scholars (e.g. Hawkins et al., 2006; Ionin et al., 2004; Liu & Gleason, 2002; Parrish, 1987; Thomas, 1989). Zegarac (2004) shared a similar view that L2 speakers of English showed a tendency of using the definite article when referents were specific and unique. The participants in the current study also exhibited a similar tendency of using *the* inaccurately for referential indefinite NPs. This finding suggests the need for L2 learners of English to understand that a specific referent requires indefinite marking when it is not identifiable to the hearer in the discourse context.

Up to this juncture, patterns of article use in the oral narratives were discussed in light of previously detailed research findings. Next, the written production of articles by the three groups is considered in comparison to the oral production of article forms. In this regard, it is important to note that limit was not placed on the amount of time participants spent on completing the task. In other words, the learners were allowed to spend as much time as they needed to complete the task. Considering the availability of time and resources for reflection and revision of the written production, an L2 writer may be able to avoid making article errors as long as the writer possesses sufficient knowledge of the English article system. If the writer lacks awareness and understanding of how articles are used in English, however, the affordance of time and the advantage of the written discourse may not be useful.

One general pattern that emerged from the analyses of the written corpus was that there was a positive correlation between the accuracy in article use and levels of English language proficiency, a finding also observed in the case of the oral narratives. The increased accuracy in the use of grammatical features along the proficiency continuum is a generally observed phenomenon in second language acquisition. Likewise, learners of English seem to develop higher levels of understanding of articles as their proficiency increases.

Between the two discourses, articles tended to be produced with higher rates of accuracy in the written discourse than in the spoken discourse. As discussed in Chapter I, this finding was expected. With more time and cognitive resources available for language monitoring and revisions, it was hypothesized that participants would produce articles more accurately in writing than in speaking. Nevertheless, distinct patterns of article use in written narratives emerged according to proficiency levels, which are discussed in the following segment.

In the beginner group, we noted that beginners still produced both article forms (*the* and

a) with low rates of accuracy in written discourse. One possible reason is that the beginners may not have sufficient awareness of the English article system to discern contexts for articles and produce them accurately. According to the Noticing Hypothesis proposed by Schmidt (1990, 2001), "learners must make conscious comparison between their own output and target language input" (Schmidt, 2010, p. 724) in order to overcome errors. This idea of "noticing the gap" originated from the scholar's own experiences in learning Portuguese in Brazil. Schmidt remarked that frequent correction for his grammatical errors had no effect in many cases because he was unaware that he was being corrected. In other words, when the learner is not aware of a particular feature of L2, this feature is less likely to be acquired or produced correctly. In this study, the beginners may have not been fully aware of the article system and thus failed to make a conscious comparison between their own writing and the target language to prevent misuses.

A second possible reason for lower levels of accuracy in the beginners' written production of articles is that L1 transfer seems to be more influential at the beginning stage of second language acquisition. In fact, L1 effect was suggested as a plausible reason for the high rates of accuracy with which beginners used certain determiners, including demonstratives and possessives, since their equivalent forms exist in Korean. Conversely, Korean does not have a morphological system that marks definiteness like English does. Thus, Korean learners of English may experience struggle in learning when and how to use articles appropriately. In fact, the beginners exhibited a strong preference for no article when they did not produce the required articles accurately. The tendency to omit articles in the written corpus as well as in the oral corpus is indicative of a strong influence of the article-less L1 on the beginners' article use.

In the intermediate group, a considerable improvement in the accurate use of *a* was observed in their written narratives. The noticeable increases in the accuracy with which the

272

intermediates employed articles in writing may be indicative of their sensitivity to rules associated with English articles. Thus, they are able to better reflect on article use. Even though the intermediates' implicit knowledge of English articles that was manifested in their oral narratives may not have reached a level of stability, they seem to have acquired a sufficient level of explicit knowledge on how articles are supposed to be used in English. As Ellis (1994) explained, explicit knowledge is what language learners use to monitor for grammatical accuracy in their language production. This kind of explicit knowledge can later become part of the implicit knowledge as they continue on their language learning.

The advanced group exhibited stable and accurate production of articles in both written and spoken discourses. The only non-target-like tendency in the advanced group was article omission, and this tendency was observed even in the written discourse. As Park (2006) revealed in his study that examined article use in essays written by Korean learners of English, article omissions seem to be the most frequent and distinct characteristics of Korean learners of English and the tendency to omit articles seems to persist even at the advanced level of proficiency. However, in other aspects of article use, the advanced group employed articles with high rates of accuracy. Thus, it can be suggested that the participants at this level of proficiency have developed an advanced level of implicit knowledge with regard to English articles as their performance in the oral and written narrative tasks attest.

In this section, patterns of article use were discussed according to proficiency groups and kinds of discourses for the purpose of addressing RQ 1 and RQ 2. Distinctive and progressive patterns of article use were observed at different levels of proficiency. It was also evident that article use was more accurate in writing than in speaking possibly due to the availability of time and attention directed for grammatical accuracy in L2 production. The patterns of article use in

oral and written narratives were indicative of the kinds of implicit knowledge that the participants have. In turn, their explicit knowledge of articles was examined by use of the think-aloud protocol, and its major findings are discussed in the next section.

6.3. Metalinguistic Knowledge of Articles

This section discusses the key findings with regard to metalinguistic understanding of articles developed by the learner groups as well as the native English speaker (NES) group. To investigate explicit knowledge of articles, the 30 participants as well as five native English speakers took a fill-in-the-article task, in which they needed to select articles (*the*, *a*, or $\emptyset$) for the blanks provided. As they were filling the blanks with appropriate articles, they were asked to verbalize reasons for their article choices in a think-aloud protocol. Reasons for article choices expressed by the participants were examined to address RQ 3 and RQ4 that are concerned with the kinds of metalinguistic hypotheses of articles exhibited by Korean learners of English at distinctive levels and their comparisons with NES participants. The results of the analyses are discussed in the context of previous research.

According to the think-aloud data produced by the participants at the beginner level of proficiency, the beginners seem to know and apply article rules that one can find in English grammar books, including the use of *the* for subsequent-mention and *a* for first-mention of singular referents. As a result, they scored relatively high rates of accuracy in instances for both articles. The kinds of hypotheses influenced by textbooks and teachers' input were classified by Butler (2002) as context-insensitive hypotheses because these learner hypotheses tend to be applied without considering relevant contextual factors. In other words, learners at lower proficiency levels in Butler's study were found to apply these rules independently without

having a clear understanding of how hearer knowledge, countability and other contextual factors are related to the article system. The patterns of article choices and reasons provided by beginners in this current study are similar to Butler's findings.

In many instances, however, textbook rules are not adequate in accounting for article use in English. For instance, Pica (1983) examined data from speech events produced by native speakers of English and compared them with commonly cited rules for article usage in textbooks. The researcher did not find the pattern of using the indefinite article for introduction and the definite article for second-mention in the corpus produced by native speakers. Thus, Pica posited that the linguistic information required for appropriate interpretation of articles is often discourse-related, which makes it difficult to master the article system through textbooks and classroom practices. Likewise, this current study found that beginners did not express hypotheses related to discourse factors and rather relied on context-insensitive hypotheses influenced by textbook rules.

When beginners made article choices, a lack of consistency was observed in the way they considered different aspects of nominals. In other words, fluctuations were found in the kinds of reasons participants provided for similar contexts. For instance, a beginner would cite 'singularity' of a referent as a reason for the choice of *a* in several instances, 'first-mention' in other identical contexts, and 'countability' in yet other contexts. Because they supplied a variety of reasons for the same article choices, I posit here that they have not developed a coherent and unified conceptualization of English articles. As mentioned above, the beginners seem to understand that factors such as first-mention, subsequent-mention, countability and noun numbers need to be considered for target-like use of articles. Still, it seems that they have not made target-like mappings between English articles and various features of nominals.

As shown in Pica (1983) and White (2010), rules provided by teachers and textbooks do not provide a coherent view of the English article system. Instead, textbooks often provide a long list of exceptions with regard to article use. For instance, White (2010) reviewed a workbook by Cole (2000) and found that the book provides 50 article rules together with 15 exceptions. The long list of rules mainly describes when to use which articles in what specific contexts. Many of these criteria for making article choices are based on collocations and particular constructions. Limitations with regard to the ways most ESL textbooks present articles may possibly explain why beginners in this current study exhibited fluctuations in the kinds of reasons for article choices and why they tended to present a variety of non-generalizable reasons when they could not apply the basic rules.

Many of the non-generalizable rules expressed by beginners were based on particular collocations and sentence constructions. For instance, several beginners in this study had developed a hypothesis that *the* is used together with prepositions such as *of* and *on*. Another hypothesis often cited by beginners was that *the* is used for nominals that refer to specific locations. These types of hypotheses were categorized in Butler (2002) as "hypotheses that show sensitivity to inappropriate contextual cues" (p. 463). The scholar explained that learners start to develop ad-hoc rules by linking article usage to inappropriate cues and tend to create certain collocation rules when they realize that the textbook rules fail to provide correct article choices. Similarly, the beginners in this study also provided ad-hoc rules based on collocations, and such a tendency was most evident in associative definite contexts.

While beginners tended to employ textbook rules inconsistently and provide non-generalizable reasons, the intermediate group was found to be equipped with more sophisticated metalanguage and comprehensive metalinguistic awareness of English articles. The percentage

of specific reasons was highest among all three learner groups, an indication of the intermediates' ability to explain article choices in an explicit manner. Also, they achieved high scores in the fill-in-the-article test, which suggests that they have achieved an acute metalinguistic awareness with regard to understanding of English articles.

Another characteristic of the intermediate group evident in the think-aloud task was the increased percentage of "combination" reasons for article choices. In other words, unlike the beginner group, the intermediates tended to consider multiple distinctions relevant to article choices. For instance, intermediates cited a combination of noun number and first-mention vs. subsequent mention distinctions as reasons for the selection of *a*. This finding is suggestive of their enhanced understanding of the multi-dimensional characteristics of the articles.

While the intermediate group exhibited a stable and accurate performance in the think-aloud task, we found that this group had difficulties in the use of *the* for associative definite nominals. That is, the majority of intermediates made non-target-like choices in associative definite contexts for kitchen-related referents (e.g., *stove*, *fridge*, *cabinet*, *timer*). The definite status of these referents can be explained by using the notion of uniqueness. Hawkins (1991) explained that *the* entails existence and uniqueness and carries a conventional implicature that there is some pragmatically shared set accessible to the speaker and the hearer within which existence and uniqueness holds. Langacker (2008b) also posited that the definite article is an indication of only one eligible candidate in the immediate discourse space. Applying this framework to the kitchen-related referents in the fill-in-the-article text, nominals such as *fridge*, *cabinet*, and *timer* are considered to exist uniquely in the shared mental space of the kitchen and thus take the definite article. The majority of intermediate participants did not have an understanding of this aspect of the definite article. Instead, they attempted to apply textbook-like

rules based on factors such as first-mention, subsequent-mention, singularity, and countability, factors that did not contribute to their selection of the appropriate article form.

A successful selection of *the* in these associative definite contexts requires an understanding that the use of articles is not a simple operation of syntax and morphology. Therefore, a rigid application of article rules would not result in the accurate selection of *the*, a tendency exhibited in the intermediate group. For instance, regardless of discourse and context factors, intermediate learners tended to employ *a* for nominals that were not mentioned before and singular. They also tended to select *the* when they find previously mentioned referents. None of the participants cited discourse factors such as the hearer's previous knowledge, shared knowledge between the speaker and the hearer, and immediate situational knowledge as reasons for article selections. In the investigation of the ways Slavic speakers of English used articles, Ekiert (2010) argued that perspective-taking and detection of what is assumed to be known to whom may be the ultimate cause of difficulty with English articles. Similarly, in this current study, participants at lower levels of proficiency failed to recognize the importance of considering the hearer's perspective and identifiability of the referent by the hearer in the immediate discourse context.

In both the beginner and intermediate groups, countability was identified as a factor that conditioned non-target-like article choices. Several scholars have demonstrated that countability perception was difficult for L2 learners of English from article-less L1 backgrounds. For instance, Yoon (1993) conducted a countability judgment test with native speakers of English and Japanese learners of English and reported that inaccurate perception of noun countability contributed to lower accuracy in the selection articles. Also, White (2009) reported that L2 learners of English had trouble in classifying abstract nouns as countable nouns. A major

challenge with noun countability among ESL learners is the underlying assumption that countability is an inherent feature of nouns in English. For instance, the think-aloud protocol in this current study revealed the assumption by Korean learners of English that certain nouns such as *water*, *gas*, *butter* and *cheese* are always inherently uncountable. They tended to choose $\emptyset$ for these nouns almost automatically without considering other discourse factors. In English, however, the majority of uncountable nouns can also be used for countable reference. For instance, *cheese* can be used for countable reference, as in *Here is a list of chesses from France*. Therefore, the assumption of treating countability as a fixed attribute of nouns may lead to non-target-like article choices.

We now turn to metalinguistic hypotheses expressed by more advanced participants. In the advanced group, the percentage of non-specific reasons was higher than the other proficiency groups, and articles chosen with non-specific reasons were mostly accurate. These results suggest that the advanced participants utilized implicit knowledge more frequently than explicit knowledge, thus not being able to verbalize reasons for article selections. A possible explanation for this tendency is related to the kinds of language learning these participants have been engaged in. One of the recruitment criteria for the advanced group was for L2 speakers of English to have lived in the U.S. for more than five years. The majority of the advanced participants attended high school in the U.S. As such, they were immersed in the target language environment and exposed to abundant language. Naturally, their language learning was likely to be more implicit than explicit. The outcome of such implicit learning would be the development of implicit understanding of article usage, which may not be readily available for verbalization. Furthermore, the advanced participants have had relatively less exposure to ESL textbooks than the other proficiency groups because the majority of them attended high school in the U.S. and

learned English in natural school settings. Lower level participants have been influenced by various ESL textbooks and learning materials, which provide them with metalinguistic terminology about language use. As such, they were more capable of discussing their article choices more explicitly, using grammatical terminology. On the other hand, the majority of advanced participants had less exposure to explicit English grammar instruction. Thus, they were not able to verbalize reasons for article choices in the think-aloud protocol as fluently as the other proficiency groups.

In turn, specific reasons provided by the advanced group for their article selections can be examined according to target articles and NP contexts. When advanced speakers selected *the* in anaphoric definite contexts, they accurately cited subsequent-mention as a primary reason. In associative definite contexts, a variety of discourse factors needed to be considered. In these instances, advanced participants could not provide specific reasons but resorted to non-specific reasons based on intuitive judgments and correctly selected *the*. These observations suggest that the advanced participants' understanding of the definite article has approximated native-like usage.

In indefinite contexts, advanced speakers tended to focus on singularity of referents as a prominent reason for the choice of *a*. With regard to the indefinite article, Christophersen (1939) posited that *a* is a unity-denoting article and this article only indicates that the referent is understood and expressed by the speaker as a unit entity. The finding that advanced speakers identified singularity of nominals as a main reason for the selection of *a* is indicative of their understanding of *a* as an article that denotes an individualized entity with a clear boundary.

Finally, the NES group, which consisted of five native English speakers with Korean heritage, seems to have employed intuitive judgments for the majority of article selections. When

NES participants provided specific reasons for article selections, subsequent-mention and specificity of referents were cited as reasons for the selection of *the*. For the indefinite article, non-specificity, singularity and first-mention were cited as factors that contributed to the choice of *a*. However, the most frequent responses for both articles were non-specific reasons in the NES group. The observation that the NES participants were not able to explicitly explain their article choices in the majority of the instances suggests that the NES group possess implicit knowledge of articles that are not easily brought into consciousness for explicit explanation. Further, these participants have acquired English as their native language and thus may not be familiar with grammatical terminology that L2 learners are exposed to in English textbooks and classroom instructions. As such, the NES participants had difficulty in explaining their article choices: They just knew intuitively what articles to use in the given contexts.

This section discussed key findings with regard to metalinguistic hypotheses expressed by the participants at various levels of proficiency. Butler (2002) pointed out that hypothesis patterns are not always clear-cut because learners with different levels of proficiency seem to mix various types of hypotheses. In the case of this study, participants employed various kinds of hypotheses for their article selections and these patterns may not be neatly divided along the proficiency groups. After all, proficiency levels are not distinct but are part of the continuum of development. Also, fluctuations in language use and constant revisions of hypotheses about various language features are a natural phenomenon in L2 language acquisition. Nevertheless, we can find that participants at lower levels of proficiency depended more on textbook-based rules and applied them in a way that was insensitive to contextual factors. As proficiency levels increased, more target-like article choices were made based on intuitive judgments by applying more implicit knowledge of articles. These observations suggest that conceptual changes with

regard to the use of English articles occur gradually along with the increase in the amount of exposure to authentic language use and the increase in the overall proficiency.

6.4. Comparison of Implicit and Explicit Knowledge of Articles

In order to respond to RQ 5 that focused on how explicit understanding of articles was manifested in the actual oral production of articles, three participants (i.e., one beginner, one intermediate, and one advanced) were selected from the three proficiency groups to serve as case studies. In each case, the patterns of articles produced by the participant in the oral narratives were compared with his/her metalinguistic hypotheses about articles obtained from the think-aloud protocol. The results of the comparative analyses revealed differences and similarities between the explicit and implicit knowledge of English articles at distinct levels of proficiency.

In the case of the beginner participant, her explicit knowledge of articles was constructed mainly based on de-contextualized textbook rules (e.g., *a* for first-mention and *the* for subsequent-mention). Target-like perception of noun countability and identification of associative definite contexts were found to be challenging to the beginner. With regard to the implicit knowledge, articles were produced accurately in the oral narrative primarily when they were used as part of memorized expressions. In other instances, articles were omitted frequently.

The intermediate participant exhibited a more sophisticated explicit knowledge of articles and achieved a good performance on the fill-in-the-article test. Nevertheless, this participant's perception of noun countability was still not target-like in several instances. Also, the application of hypotheses without considering contextual features resulted in incorrect article choices in associative definite contexts. For instance, the intermediate learner did not seem to consider the perspective and prior knowledge of the hearer with respect to referents, a factor

essential for the accurate selection of articles.

In turn, patterns of article production in the oral narrative of the intermediate speaker revealed fluctuations in the use and omission of articles. While the participant scored high accuracy for both articles in the think-aloud protocol, he produced *a* with much lower rates of accuracy than *the* in his oral narrative. This finding is indicative of how the accurate production of the indefinite article was considerably affected by cognitive processing constraints in a spontaneous speech at the intermediate level of proficiency. In addition, the observed fluctuations in the use of articles in speaking are suggestive of how the participant's implicit knowledge is developing and changing towards target grammar.

At the advanced level of proficiency and its relation to implicitness, we found that the advanced participant was not familiar with meta-language and thus was not able to articulate the reasons for article choices. Many article choices were correctly made based on intuitive judgments, an indication that the participant was applying his implicit knowledge of articles in carrying out the fill-in-the-article test. In the oral narrative, *the* was produced with high level of accuracy, but the accuracy in indefinite contexts was lower mainly due to the tendency to over-use *the* in indefinite instances. In short, stable and accurate performances in both the oral narrative task and the think-aloud task suggest that the increased language use and exposure to various uses of articles in diverse contexts may help L2 learners of English achieve a near-native competence of English articles.

In the comparisons between implicit and explicit knowledge of English articles, certain differences were found in the three cases with regard to the way learners used articles and the way they explained article use in explicit terms. In general, the beginner participant and the intermediate participant tend to be influenced more by textbook rules and explicit instructions

283

they have received about article use. Since textbook rules provide discrete usage types for different articles, these rules are not effective in assisting L2 learners of English to develop coherent and unified conceptualizations of articles. In particular, the intermediate participant was able to recite textbook rules about articles relatively accurately and made article choices with high accuracy. Yet, actual production of articles in the oral narrative did not live up to the level of explicit knowledge. This result suggests that the knowledge of articles remains explicit in the mind of the intermediate learner and has not caused conceptual restructurings of nominal expressions in English.

On the other hand, the advanced participant exhibited stable performance in both think-aloud and oral production tasks. Thus, it seems that the advanced participant exhibits stability in his implicit knowledge of articles. Here, it is important to note that the performance of the advanced participant is still lower than that of NES participants. The advanced level of implicit knowledge of articles evidenced in the advanced participant may be attributed to the implicit nature of language acquisition experiences. The sociolinguistic interview revealed that the participant has learned English implicitly in an immersion environment since he moved to the U.S. with his parents at the age of 15. With constant exposure and input to authentic language, the participant developed an implicit understanding of article use without much influence of grammar books and explicit instructions.

Up to this juncture, key findings were discussed in light of previous research as they related to the five research questions posed for the current study. In the next section, theoretical implications of the current study are discussed with regard to L2 acquisition of articles.

284

6.5. Theoretical implications

In this section, theoretical implications of the study are considered with regard to second language acquisition of English articles. First, Berezowski's position for the two-member article system is discussed in the context of the current study. Then, a new approach to understanding articles is suggested based on the tenets of Cognitive Grammar, which suggests that linguistic units serve as prompts for conceptual operations and that meaning construction is a dynamic process. Thus, the suggested approach will focus on how conceptual operations involved in the use of the indefinite article are different from those for the definite article.

In Chapter II, Berezowski's (2009) incomplete grammaticalization model was discussed in detail to explain why the current study assumes the existence of only two overt article forms (*the* and *a*). According to this model, articles in English have been going through the process of grammaticalization. Originally, *the* evolved from a demonstrative, and *a* from a numeral. Over time, these forms have been used in more extensive grammatical environments. For instance, as Langacker (2008b) describes, *the* gradually lost its proximal/distal distinction as a demonstrative and began to be used in instances in which there is only one evident referent identifiable to both the speaker and the hearer. In the case of the indefinite article, *a* gradually expanded its use from representing number *one* to a more extensive reference to nominal entities that are considered as bounded singular units. Further, the expansion of grammatical environments for article use is still incomplete. Therefore, Berezowski argues that instances in which an overt article is not employed are a direct consequence of incomplete article grammaticalization. Because these are gaps in the use of articles in English, a covert article does not need to be posited.

As this study was informed by Berezowski's (2009) theory, the analyses of the data and discussions of the findings focused on the use and acquisition of the two articles. Nevertheless,

in order to ensure a complete examination of article use and acquisition by Korean learners of English and to reflect the instructional reality that most ESL / EFL textbooks assume the existence of the zero article, the current study also examined how the participants dealt with instances in which no article is required. Quantitative analyses on the use of articles in the oral and written narratives revealed that participants in all proficiency groups scored high rates of accuracy in contexts in which articles were not required. Recall that even the beginners exhibited high accuracy in instances in which no article was required in both discourses. This observation suggests that bare NP contexts do not pose considerable difficulties to Korean learners of English.

We also note that participants did not refer to the term "the zero article" in the explanations for article choices during the think-aloud protocol. Rather, their remarks were along the lines of "no article is necessary" and "nothing is needed here" when no overt article form was required. These comments suggest that the reality of the target language perceived by the participants as well as native English speakers points to English having two article forms.

Considering that the origins and the uses of the two articles are distinctly different, we may need to adopt different theoretical approaches to understanding the two articles. In this regard, it is worth reviewing Christophersen's (1939) discussion of English articles. Christophersen (1939) posits two distinct semantic attributes to the two articles. Specifically, the notion of familiarity is used to describe the meaning of the definite article. The scholar notes, "a condition for the use of *the* is that there is a basis of understanding between speaker and hearer" (p. 28). In other words, *the* can be used when there is a mutual familiarity with the referent.

In turn, *a* is described by Christophersen (1939) as a morpheme that denotes unity. As such, the indefinite article turns an abstract and amorphous notion into a unit word. Similarly, White (2010) describes *a* in terms of *individuation*. For instance, *cake* becomes a unit word when

it occurs with the unity-denoting indefinite article in reference to a particular cake, as in *I bought a cake for the birthday party*. When *cake* is used without any article, as in *I want to eat cake*, it is understood as having no precise limit. Christophersen (1939) points out that *a* is neutral with regard to familiarity. In other words, the unity-denoting article *a* does not concern whether the speaker and the hearer are familiar with the referent. It only indicates that the referent is understood and expressed by the speaker as an individuated entity.

Christophersen's observations have important implications regarding how we theorize second language acquisition of articles. In the following segments, the two articles are discussed separately, and these discussions are couched within Cognitive Grammar. As stated in Robinson and Ellis (2008), "language has come to represent the world as we know it; it is grounded in our perceptual experience" (p. 3). In other words, linguistic forms are used as symbolic tools to represent the language users' points of view about the world around them. From the cognitive linguistics view of language, conceptualization is crucial to understanding language. Therefore, the following discussion of the two articles is predicated on the notion of conceptualization. Conceptual diagrams of the two articles will be first presented followed by an integrative discussion of second language article acquisition which centers on how nominals are conceptualized and expressed in English.

First, the use of the indefinite article signifies that the speaker conceptualizes a nominal referent as a bounded unit entity. By employing *a*, the speaker imposes on the hearer the interpretation of the referent as a singular entity. Since the notion of countability is closely related to the indefinite article in English, it is necessary to review what a countable noun means. Langacker (2008b) defines a countable noun as "being bounded within the immediate scope in the domain of instantiation" (p. 132). In other words, a nominal entity is deemed countable when

it is perceived as having a precise limit in the immediate discourse context. Categorization of

countability depends on how things are conceptualized in one language, which to some extent is

independent of their objective nature. For instance, *tree* and *wood* may have the same material

substance, but the first is considered countable and the latter, uncountable. Further, countability

is not an inherent and fixed attribute of a lexicon but determined by the context and the speaker's

intent. Thus, the same nominal can be conceptualized differently with respect to countability. The

use of *a* for a singular referent can be visually represented in contrast to an uncountable referent

and a plural form of a countable referent, as in Figure 6.1. These diagrams are modified from

Langacker (2008b).

Figure 6. 1. Diagrams of a singular referent, an uncountable referent and a plural referent
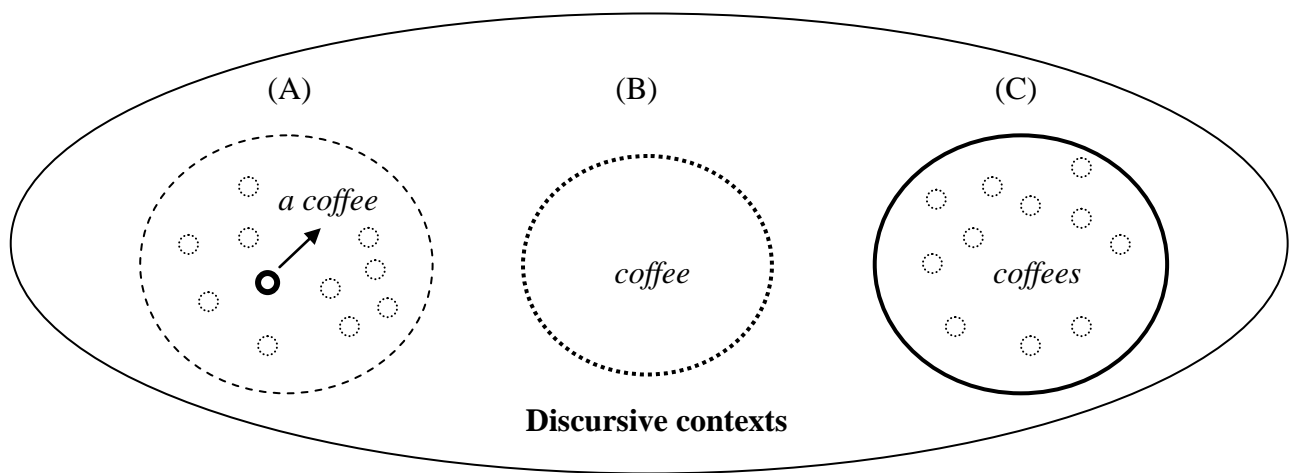
The above figure provides visual representations of a singular referent, an uncountable

referent, and a plural referent and the context-dependent nature of countability conceptualization.

For instance, in (A), *coffee* is conceptualized as an entity with a precise boundary, as in *I had a

coffee and a bagel with cream cheese for lunch*. The lexical item *coffee* can have many potential

instantiations as represented by dotted circles, and *a coffee* represents one instantiation of the

lexicon. When individual singular entities are conceptualized together as a group in the speaker's

mind in a given context, it can be expressed as a plural form, *coffees*, as in *I ordered three coffees and a hot chocolate*, and it can be visually represented as in (C). Meanwhile, *coffee* can be conceptualized as unbounded and uncountable when the entity focuses on its substance rather than imposing a precise boundary, as in *Coffee causes fatigue, exhaustion, and addiction*. Therefore, the uncountable reference can be represented like (B). Here, the dotted boundary illustrates that a limit is not imposed on *coffee*, and it can expand in any shape or size conceptually. The speaker may choose to use one of the three constructions of nominals depending on how he/she wishes to present the nominal to the hearer in a particular discursive context as represented by the outer layer in the diagram.

While most ESL textbooks tend to present the definite article and the indefinite article as two forms in semantic opposition, as Ekiert (2010) points out, these two article forms are not in complementary distribution. Specifically, the indefinite article is related to the domain of countability as discussed in the earlier segment. In turn, *the* can appear before all the three kinds of referents (i.e., a singular, countable referent, an uncountable referent, and a plural referent) when certain semantic, pragmatic, and discursive conditions are satisfied. In other words, the need for the definite article in English is dependent on multiple layers of semantic, pragmatic, and discourse factors that predicate in English. These layers can be represented as the following.

Figure 6. 2. Layers of factors that influence the use of the definite article in English
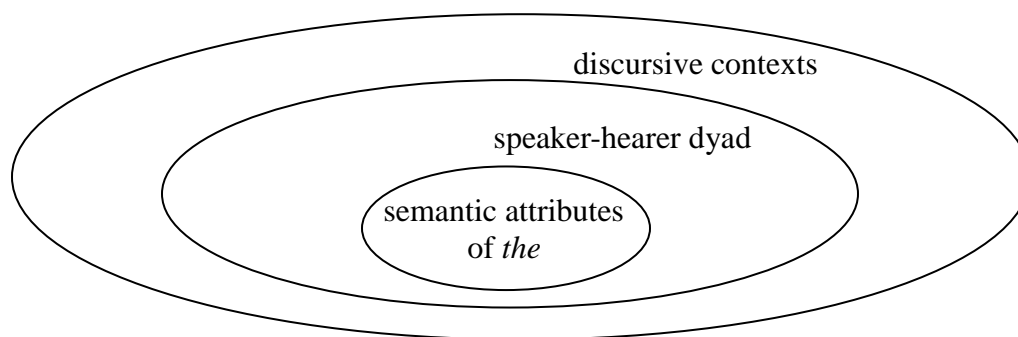
The above figure is illustrative of three major layers of variables that influence the use of the definite article in English. At the core are basic semantic attributes of *the*, which are related to the concept of definiteness. As discussed in detail in Chapter II, definiteness has been the focus of scholarly interest for more than a century. The definite article is a semantic marker that indicates that there is only one eligible referent that is mutually familiar and identifiable.

The notion of familiarity and identifiability is a relative term, and the concern is to whom it should be familiar in order to warrant the use of the definite article. This factor is represented as the "speaker-hearer dyad" layer in Figure 6.2. In English, the definite article is used when a referent is assumed to be familiar to both the speaker and the hearer. Therefore, the speaker needs to constantly assess what is assumed to be known to the hearer by considering the perspective of the hearer. A referent that the speaker wishes to mention may take the definite article in one instance but not in another instance. For instance, one may say to his son, "*let's go to the park to throw some baseball*" because the hearer is assumed to know which park the speaker is referring to. Yet, the same person may say to a colleague at work, "*I went to a park yesterday to throw some baseball with my son*" when the speaker thinks that the hearer would not be familiar with the park he often visits.

The final layer is the discursive context layer that also influences the use of the definite article. Depending on the discourse context, certain referents may take the definite article. For instance, nominal referents that are physically present in the immediate space can be used with *the* even though they have not been previously mentioned. A larger cultural and social context may also play a role in the use of the definite article.

Up to this juncture, the two articles were discussed separately using diagrams to illustrate how articles are conceptualized in English. With regard to article use, countability and

definiteness can be conceived as attributes of nominals in English which are highly context-dependent. Also, these attributes are not determined in a sequential manner but concurrently in the mind of the speaker. For example, depending on the speaker's intent on expressing a nominal in a particular context, he/she may choose to say *the coffee* as opposed to *coffee*, *a coffee*, *coffees*, and *the coffees*. In doing so, the speaker selects a nominal phrase that best describes the referent that he/she wishes to communicate to the hearer. Given these characteristics, I suggest that variability in L2 article use need to be discussed in an integrated manner by placing nominal conceptualizations at the center of the discussion.

The definite and indefinite articles in English are concerned with the nominal attributes of countability and definiteness, which are conditioned by speaker-hearer relationship as well as the discursive factors. Thus, patterns of article use and errors by L2 learners of English need to be approached with respect to how nominals are conceptualized in a discourse. In other words, when learners conceptualize nominals in a non-target-like manner in a given context, they may produce articles inappropriately.

In particular, when conceptual categories with regard to nominal expressions differ between the L1 and L2, learners may produce articles inaccurately due to conceptual transfer. For instance, there are cross-linguistic differences in how nominal entities are conceptualized with regard to countability. In other words, certain nominals may be conceptualized as bounded entities with precise limits in one language, but their equivalent forms in another language may be conceptualized as uncountable. For instance, as Park (2006) points out, *homework* is generally considered uncountable in English. If the speaker wants to count homework, he would use *many homework assignments* instead of **many homeworks*. In Korean, however, 숙제 (homework) is

perceived as a countable noun. These cross-linguistic differences exist because different languages may carve up events and experience in different ways. As Yoon's (1993) study on the perception of countability of English nouns attested, these differences present obstacles to L2 learners in English in their attempts to acquire the use of *a*.

Another source of difficulty in the way L2 learners conceptualize nominals with regard to countability is that the learners may have been influenced by the traditional grammar instructions that treat mass nouns and countable nouns as fixed categories of nominals. As several participants suggested in the think-aloud protocol, they seem to have a notion that certain nouns are always countable and other nouns are always uncountable without reference to contextual factors and the speaker-hearer relationship. The fixed dichotomy of nouns based on countability may contribute to non-target-like article selections.

With regard to the definite article, the complexity involved in the expression of definiteness in English presents a considerable challenge for L2 learners of English, and especially when their L1s do not formally express definiteness in a similar way as English does. As Robertson (2000) explained, languages that do not have an article system such as Korean and Chinese do not formally and morphologically mark definite status of a nominal reference. As such, information about definiteness is mostly retrieved by the hearer implicitly based on contextual cues. Therefore, if Korean learners of English have not fully internalized the conceptualization of nominals with regard to definiteness, they may produce articles mainly based on countability of nominal reference.

Furthermore, in the current study, the think-aloud results suggested that participants tended not to consider the perspective of the reader and the contextual factors in making article choices. This tendency may be attributed to traditional grammar instructions that present simple

article rules such as using *a* for first-mention and *the* for second-mention. Thus, participants at lower proficiency levels had difficulty in selecting *the* for kitchen-related referents when the definiteness of these referents was influenced by discursive context.

In this section, theoretical implications of the current study were discussed with regard to second language acquisition of English articles. First, the two articles were discussed separately as they pertain to nominal attributes of countability and definiteness respectively. Then, the two articles were discussed together from the perspective of L2 acquisition of articles. Several reasons for non-target-like use of articles by L2 learners of English were suggested as they relate to non-target-like conceptualizations of nominals with respect to countability and definiteness. The influence of traditional article instructions, conceptual transfer between the L1 and L2, and the interplay between the two attributes of nominals are indicative of a need for a different pedagogical approach to L2 article instruction. In this respect, the next section discusses pedagogical implications of the current study and offers suggestions for article instruction.

6.6. Pedagogical Implications

The motivation for this current study was the perceived difficulties in the acquisition of English articles by L2 learners of English, particularly when the learners' L1s utilize different ways of conceptualizing and expressing definiteness of nominal entities. Challenges with regard to the L2 acquisition and use of English articles have been investigated for many decades (Butler, 2002; Ekiert, 2010; Huebner, 1979; Park, 2006, Young, 1996, among others) and several pedagogical suggestions have been made (e.g., Akakura, 2009; Master, 1990, 1997, 2002; Pica, 1983). Nevertheless, articles remain a challenge to L2 English language learners and teachers alike. As this current study examined how L2 learners of English from Korea understand and

produce articles in a comprehensive manner, the findings of this study have important

pedagogical implications. This section discusses the results of this study that are relevant to what

ESL / EFL teachers can expect about L2 learners' knowledge and the use of articles at different

levels of proficiency and what pedagogical approaches can be taken for effective article

instruction.

The results of this current study can help ESL / EFL teachers develop a better

understanding of what their students may experience with regard to article acquisition. In turn,

this understanding can help teachers design a more effective instruction of English articles.

Before discussing pedagogical implications of characteristics in article use at different

proficiency levels, it is important to note that the distinction of proficiency levels used in the

current study may not necessarily respond to proficiency levels used in other instructional

settings. Therefore, caution is advised in interpreting the findings of the current study with regard

to proficiency levels.

First of all, the study was able to attest that article use becomes more accurate as the

overall language proficiency increases. In other words, teachers may notice frequent article

errors and non-target-like use of articles in both oral and written discourses produced by

beginner-level learners. As their overall language proficiency increases, articles tend to be

produced with higher rates of accuracy. Yet, there may still be a persistent tendency to omit

articles, and this tendency may be more evident among learners whose L1s do not have articles.

Secondly, between the two article forms, the study also revealed that L2 learners of

English generally employed *the* more accurately than *a*. Since *the* seems to be acquired before *a*

(e.g., Huebner, 1979; Parrish, 1987), teachers may focus on the use of the definite article for

beginning learners while delaying an extensive discussion of the use of the indefinite article until

the learner's proficiency reaches a higher level. Indeed, the current study noted a considerable improvement in the accuracy with which *a* was produced in the written discourse of the intermediate group. This observation suggests that intermediate learners may be at the threshold of acquiring appropriate uses of English articles.

Thirdly, with respect to the use of the definite article by L2 learners, teachers may note that non-target-like production of *the* is often attributed to a learner hypothesis which suggests that the definite article is used for a specific nominal referent. In fact, this hypothesis may not only be a learner hypothesis but also a hypothesis upheld by ESL / EFL teachers. For instance, White (2010) found that the term 'specific' was used extensively by ESL teachers to describe certain uses of *the*. In his study, 12 experienced ESL teachers were asked to provide explanations for 20 instances of articles. The results showed that 'specific' was the most frequently used term for the participating teachers to explain various uses of *the*, followed by 'definite' and 'collocation'. When teachers continue to use the term 'specific' to account for the use of *the*, learners may develop a hypothesis that *the* is needed to refer to a specific noun. However, teachers need to understand that *the* is required when a nominal refers to an entity that both the speaker and the hearer recognize as the only eligible referent in the discourse, not when a nominal refers to a specific entity.

Fourthly, teachers may also keep in mind that a multiple factors influence the non-target-like production of *a* by L2 learners of English. The first factor is the consideration of the hearer's perspective. As mentioned earlier, a specific referent may take the definite article when the referent is already present in the immediate discourse space shared by the speaker and the hearer. Conversely, the indefinite article is required when it refers to a singular referent that is specific in the mind of the speaker but is not yet known to the hearer. Thus, the speaker needs to consider

the perspective of the hearer before selecting which article to use. Indeed, Ekiert (2010) claims that the assessment of what is known to the hearer may be a prime suspect for the difficulty that English articles pose for Slavic learners of English. Therefore, teachers need to incorporate the notion of familiarity of a referent to the hearer in their instruction of articles.

Another factor that influences errors in the use of *a* is the difficulty in determining countability of nominal referents. In English, the indefinite article is employed for a noun that is countable, singular, and unknown to the hearer. This current study suggested that Korean learners of English often misjudged countability of NPs, resulting in non-target-like article choices. As explained in the previous section, this factor may be more relevant to English language learners whose L1s have different ways of conceptualizing countability of nouns. Teachers can be informed of this observation and find a way to help learners develop an appropriate understanding of countability in English. It is also important for teachers to remind learners that countability is not an inherent feature of a lexicon but is dependent on a particular context.

Finally, with regard to the explicit understanding of articles and its relations with the actual use of articles, this current study suggests that there may be gaps between the explicit and implicit knowledge of the article system at lower proficiency levels. For instance, even if learners achieve good scores on explicit tests of articles (e.g., fill-in-the-article test or multiple-choice questions), they may use articles with low rates of accuracy in discourses that are largely spontaneous. In other words, learners may be able to recite article rules that they have learned from textbooks and other instructional resources, but their explicit awareness of articles may not indicate a similar level of implicit and unconscious ability to produce articles accurately. Conversely, an inaccurate use of articles does not necessarily suggest that learners do not have an understanding of articles. Teachers may note that inappropriate article use may not be an

indication of the absence or lack of knowledge of articles but article errors can be conditioned by cognitive constraints with regard to pressure to perform in L2. In this regard, in order to help bridge the gap between explicit and implicit knowledge, teachers may design classroom activities that encourage learners to practice using articles in various discourse contexts. For instance, as Pica (1983) suggested, activities such as ordering food at restaurants, asking for directions on the street, and making dinner in the kitchen can provide opportunities for learners to develop an increased sense of countability and definiteness as they relate to article choices in English.

The above segment discussed learners' use and understanding of articles that the current study found and their pedagogical implications. We now turn to the perspective of English language teachers. In general, ESL / EFL teachers have not been well equipped with effective article instruction strategies. First of all, ESL teachers who are native speakers of English may find it difficult to explain various uses of articles in an explicit and coherent manner to their students. In White's (2010) survey, many experienced ESL teachers expressed a lack of confidence and a sense of frustration with regard to teaching articles. One participant who has been teaching ESL for 23 years confided, "I feel that after a certain point, it is all intuition. Even highly proficient L2s can't get it, and it gets frustrating both for the L2 learner and for me. I've given up on teaching articles a long time ago and usually only address it when a particular student really wants to discuss it" (p. 50). This teacher's remarks resonate with how the NES participants in this current study relied on intuitive judgments in making articles choices for the think-aloud task. Thus, native English speaking teachers may not know how to explain article use to their students when their own article use is based on unconscious explicit knowledge and intuition.

When teachers cannot come up with satisfactory explanations for various kinds of article use, they may turn to ESL textbooks and other resources. However, as Pica (1983) and White (2010) demonstrated, most ESL textbooks only provide a list of article rules based on collocations and expressions rather than offering a coherent and comprehensive framework on article use in English. Thus, one ESL teacher in White's (2010) survey stated, "I don't feel very comfortable teaching articles mainly because many uses of articles don't follow any rules introduced in grammar books" (p. 49). This concern is more problematic for EFL teachers who are mostly non-native English speakers themselves. As they tend to rely more on textbooks, article instruction in EFL environments may reinforce the idea to the learners that articles are employed based on a list of fixed usage rules.

Differentiated pedagogical strategies can be suggested for different proficiency levels based on the findings of the current study. For lower level learners of English, learners may benefit from learning useful formulaic sequences[15] since these expressions tend to contain various instances in which articles are used. As Gor and Long (2009) point out, L2 learners exhibit a tendency to rely on expressions that they have memorized when placed under pressure to produce stories in L2. The current study also found that beginners were able to produce articles relatively accurately when the article forms were used as part of familiar expressions that the learners have memorized as a whole. When lower level learners expand a pool of useful formulaic sequences they can draw from in spontaneous L2 production, their articles can be provided more accurately, and in turn, they may begin to acquire articles in an implicit manner.

---

[15] Wray & Perkins (2000) defines the formulaic sequence as "a sequence, continuous or discontinuous, of words or other meaning elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar" (p. 1).

In this respect, the usefulness of learning formulaic expressions for article acquisition can be understood as part of the usage-based approach. Langacker (2008a) explained that "native speakers control an immense inventory of conventional expressions and patterns of expression enabling them to handle a continuous flow of rapid speech" (p. 84). In other words, building a rich inventory of formulaic expressions that contain various uses of articles will help less proficient learners of English enhance fluency and accuracy in L2 language production.

In the classroom, teachers can design activities and tasks that can naturally expose learners to situations in which various article uses are needed. For instance, Pica (1983) proposed that activities such as ordering food at restaurants and asking for directions can be non-frustrating lessons for beginner learners. Prior to a classroom activity, sample dialogues can be provided to show how articles are used for various referential expressions. Then, learners can create similar dialogues on their own while carrying out communicative tasks. In this process, they can practice and acquire useful conventionalized expressions and chunks.

For more advanced learners, explicit instruction may be helpful. For instance, Akakura (2009) used a Web-based instructional program that is designed based on a cognitive approach to articles for the purpose of examining the effect of explicit instruction on implicit knowledge of English articles. The study revealed that the participants achieved gains on article accuracy in both the explicit task and the oral production task after receiving the explicit instruction. This finding suggests that explicit instruction plays a role in promoting the implicit knowledge of articles.

When teachers give explicit instructions on English articles, they need to stay away from the traditional approaches to article pedagogy that place the definite article and the indefinite article in complementary distribution and provide a list of usage patterns for each article. Instead,

299

teachers need to focus on explaining how articles are dependent on nominal conceptualizations, speaker-hearer relationships, and discursive contextual factors. For instance, countability is an attribute of a nominal that can be conceptualized differently in accordance with the speaker's intention as well as the discursive context. Traditionally, grammar books provided a list of mass nouns and a separate list of countable nouns. However, teachers can explain to the learners how the same noun can be constructed as countable or uncountable depending on its reference and the context. For instance, *coffee* can be used as a singular form as in <u>*A coffee* *a day will keep your*</u> *work pains away* to refer to a cup of coffee, and a plural form as in *The Thai coffee industry* *needs to shift its focus to specialty <u>coffees</u>* to indicate different kinds of coffee beans. Advanced learners can recognize that countability is not an inherent and fixed attribute of a nominal but is a context-dependent property.

Similarly, definiteness is a semantic and pragmatic attribute of a nominal which is highly dependent on contextual factors. As discussed in detail in Section 6.5, *the* can be used as part of a nominal phrase when the nominal is conceptualized as something familiar and identifiable to both the speaker and the hearer. Teachers need to emphasize the context-dependent characteristics of the definite article so that learners can become sensitive observers of article use in various discursive settings. In the classroom, the visual representations of the two articles introduced in Figures 6.1 and 6.2 can be used. These visual diagrams will be useful as an overall guide in analyzing various article uses that more proficient learners are exposed to.

Ultimately, the most important aspect of L2 article acquisition is constant exposure to authentic article use with a heightened awareness and sensitivity to articles used in various contexts. According to Cognitive Grammar (Langacker, 2008a), a language is comprised of an inventory of meaningful elements that have been conventionally established and shared by its

speakers. These conventional elements are linguistic units that have been abstracted from particular usage events. Therefore, Langacker (2008a) explains that mastering a language requires specific, usage-based learning of a variety of conventional units and suggests the importance of providing L2 learners with sufficient exposure to representative uses of a given linguistic unit. When this explanation is applied to L2 acquisition of articles, it can be suggested that continued exposure to various uses of articles in a wide range of discourse contexts is essential for learners to develop target-like conceptualizations of these forms.

Chapter VII. Conclusion

As outlined in Chapter I, the current study aimed to expand our understanding of the challenges associated with L2 acquisition of the article system in English by conducting a comprehensive investigation of how Korean L2 speakers of English understand and employ English articles. The study revealed that patterns of article use varied not only according to the external variable of overall language proficiency but also with respect to several linguistic variables. For instance, participants were found to omit articles more frequently when nominal referents were preceded by modifying adjectives. Also, the variable of categories of NP contexts seems to influence article use (i.e., anaphoric definite and associative definite contexts for *the*, and referential indefinite and non-referential indefinite contexts for *a*). In addition, the manner with which articles were employed by the participants varied between the oral and the written discourses. Since writing is a planned discourse, articles were generally produced with higher rates of accuracy in written narratives than in oral narratives.

The current study also gathered qualitative data on the explicit understanding of articles by using a think-aloud protocol. In other words, the participants verbally provided reasons for

their article choices as they were completing the fill-in-the-article test. The results showed that

some instances, the beginners tended to apply explicit article rules that are commonly presented

in textbooks but in other instances, they applied non-generalizable rules that were mostly based

on certain collocations. The intermediate participants exhibited a similar tendency to rely on

certain article rules and selected articles with a high accuracy. However, in the majority of

associative definite contexts, they failed to recognize the hearer's perspective and the larger

contextual factors and selected inaccurate articles. Meanwhile, the advanced participants as well

as the NES participants tended to exercise intuition in making article choices, and this tendency

was considerably more evident in the NES group than the advanced group. These observations

suggest that there is a qualitative difference between the beginner and intermediate participants

on one side and the advanced and NES participants on the other side. With the increase in

proficiency and exposure to article use in a variety of authentic contexts, L2 learners of English

may gradually acquire how nominal entities are conceptualized in English with regard to

countability and how their definite status is determined by multiple variables that are semantic,

pragmatic, and contextual in nature.

In concluding the report of the current study, limitations of the study are discussed and

suggestions for future research are presented in the following sections.

7.1. Limitations of the Study

There are several limitations to this current study. The first of these are the limitation

that has to do with the grammaticality judgment task employed in the study. As discussed in

Chapter V, the grammaticality judgment task yielded mixed results. Specifically, there was no

statistically significant difference between the performance of the beginner group and that of the

advanced group. Several reasons can be proposed for this outcome. The first possible reason is that participants might have employed strategies in making grammaticality judgments that were not relevant to knowledge of English articles. Indeed, Ellis (1991) reported that (1) the Chinese learners of English in his study demonstrated considerable inconsistency in their judgments, (2) learners employed a variety of strategies in making judgments, including feeling, rehearsing, using analogy, evaluating a sentence, and guessing, and (3) learners were relatively unsuccessful in using explicit grammatical knowledge to make judgments. It is plausible that the participants in the current study also adopted various test-taking strategies instead of using their knowledge of articles.

The second concern with regard to the grammaticality judgment task is that the test items might have been unchallenging to the participants in this current study. Most of the dialogues used for grammaticality judgment were adopted from Ionin et al., (2004), whose test items were constructed to examine fluctuations of article choices with regard to specificity of noun referents. As the current study had different objectives, new test items could have been designed for the grammaticality judgment task to obtain more meaningful results. In addition, there is a possibility that the think-aloud task provided a priming effect on the performance of the grammaticality judgment task. Prior to making grammatical judgments about various instances of article use, the participants had engaged intensively with the grammatical feature of articles while selecting articles in the blanks and verbalizing reasons for their article selections. As such, their sensitivity to English article use was heightened to a certain degree. The enhanced level of awareness of English articles may have thus influenced the result of the grammaticality judgment task, which could partly explain why the three proficiency groups performed with similar rates of accuracy.

The second limitation is concerned with the way the narrative production task was designed. In the current study, participants were allowed to narrate any stories they wanted. As such, the content of the narratives collected for data analyses varied greatly. As a wide range of nominal referents were produced in a variety of stories, a comparison of article use for the same NPs across participants was not possible. It would be interesting to use the same prompt to elicit spoken data. For example, participants may watch a short movie and then orally re-count the movie to an interlocutor. Then, there could be some level of control over the content of the oral data. This way, the same nominal referents that appear in the movie could be tracked for article use among different proficiency levels.

Lastly, the current study recruited Korean learners of English who were living in the U.S. at the time of the data collection. That is, this population has been exposed to articles used by native English speakers in various situations. The current study did not examine patterns of articles used by learners who are in Korea (i.e., EFL learners). It can be assumed that sources of input for article use in the EFL context would be mostly textbooks and a limited amount of usage input from classroom discourses. Therefore, a comparison of article use between Korean learners of English in the U.S. and Korean learners of English in Korea would provide an opportunity to investigate relationships between patterns of article use and the quality of input regarding authentic article use.

7.2. Directions for Future Research

Based on the findings of the current study, several suggestions can be made for future research. First, a study can be designed to include learners of English from L1s that have similar article systems as English in order to determine how different conceptualizations of nominal

entities with regard to definiteness and countability in different L1s have effects on L2 article use. Patterns of article use as well as metalinguistic awareness can be compared between [+ART] group and [-ART] group to identify characteristics of article use that can be attributed to L1 as opposed to more general L2 learner characteristics observe in both groups.

Secondly, a conversational task can be employed to collect data on how interactions between the two interlocutors in a discursive setting influence L2 article use. The current study elicited articles produced in narratives, which did not have an interactional element. As described in Figure 6.2, the use of article is dependent on the speaker-hearer dyad as well as relevant discursive and contextual factors. For instance, participants can work in pairs in an information-exchange task. Further, I wish to investigate how patterns of article use may be conditioned by different pairings of interlocutors (e.g., a pair of non-native English speakers from Korea as opposed to a pair consisting of a native English speaker and a Korean learner of English). In doing so, one can test a hypothesis that English language learners may use articles more accurately when they converse with native English speakers if appropriate input is provided in the immediate discursive context.

Thirdly, phonetic perception of articles by L2 learners from various L1 backgrounds can be investigated. As pointed out by Master (1997), articles are among the most frequently used words in English. Thus, learners are likely to be exposed to abundant input containing various uses of articles. However, as the current study suggested, L2 article acquisition is still challenging. One plausible factor may be related to the fact that the pronunciations of articles tend to be reduced in oral speech, making it difficult for learners to audibly perceive the articles in spoken input. In this respect, a study can be designed to test phonetic perception of articles by learners of English at different proficiency levels for the purpose of determining whether

difficulties in auditory perception of articles in the input are a factor that contributes to the challenges in L2 article acquisition.

Finally, a more longitudinal study may show how learners' conceptualizations of nominal entities with regard to English articles develop over time as their exposure to English grows and their overall language proficiency increases. As pointed out by Master (1987), "a true picture of article acquisition should be based on longitudinal studies" (p. 26). In particular, a verbal protocol such as a think-aloud protocol may be employed with the same participants over a span of time. A verbal protocol is useful in getting access to the learners' cognitive changes about English articles over time with the increased exposure to English language input and other relevant variables.

Appendix A: The Cloze Test

Passage 1

Long Lost Note

Have you ever sealed a note inside of a bottle and thrown it into the sea? Many people have done this. Some have done it for fun, and others have done it for more serious reasons.

In 1784, a young Japanese sailor threw a bottle _____ a message into the sea. The sailor had been _____ a treasure hunt in the Pacific Ocean. A storm _____ come up, and his small ship had been wrecked. _____ and the other crew members had landed on a _____ island.

At first, the men were happy because they _____ safe from the rough waves and the terrible wind. _____ waited for the storm to end. Finally the Sun _____ out. The men looked around and saw a few _____ trees lay on the ground. There was nothing to _____ except for some tiny crabs. Even worse, there was _____ water to drink.

Soon the sailor was afraid that he _____ his friends would never leave the island. They would _____ see their families again. "Still," though the sailor, "I _____ be able to send them a message."

He found _____ bottle in the wrecked ship. He cut thin pieces _____ wood from one of the trees. These pieces of _____ would serve as paper. Slowly the sailor carved the _____ about the wreck into the wood. Then he put _____ message in the bottle and sealed it well. He _____ the bottle as far as he could into the _____.

The sailor and his friends never left the island, _____ the bottle did. It rode the ocean waves for _____ years. Then one day the bottle washed up on _____ shore. A man found it tangled in some seaweed. _____ was very surprised. The bottle had handed in the _____ village that was the sailor's home. The year was _____ 1935. The sailor's message had floated at sea for 150 _____!

Passage 2

The Expression of Emotions

Joy and sadness are experienced by people in all cultures around the world, but how can we tell when other people are happy or despondent? It turns out that the expression of many emotions may be universal. Smiling is apparently a universal sign of friendliness and approval. Baring the teeth in a hostile way, as noted by Charles Darwin in the nineteenth century, may be a universal sign of anger. As the originator of the theory of evolution, Darwin believed that the universal recognition of facial expressions would have survival value. For example, facial expressions could signal the approach of enemies (or friends) in the absence of language.

Most investigators concur that certain facial expressions suggest the _____ emotions in all people. Moreover, people in diverse cultures _____ the emotions manifested by the facial expressions. In classic _____ Paul Ekman took photographs of people exhibiting the emotions _____ anger, disgust, fear, happiness, and sadness. He then asked _____ around the world to indicate what emotions were being _____ in them. Those queried ranged from European college students _____ members of the Fore, a tribe that dwells in _____ New Guinea highlands. All groups, including the Fore, who _____ almost no contact with Western culture, agreed on the _____ emotions. The Fore also displayed familiar facial expressions when _____ how they would respond if they were the characters _____ stories that called for basic emotional responses. Ekman and _____ colleagues more recently obtained similar results in a study _____ ten cultures in which participants were permitted to report _____ multiple emotions were shown by facial expressions. The participants _____ agreed on which two emotions were being shown and _____ emotion was more intense.

Psychological researchers generally recognize that _____ expressions reflect emotional states. In fact, various emotional states give _____ to certain patterns of electrical activity in the facial _____ and in the brain. The facial-feedback hypothesis argues, however, _____ the causal relationship between emotions and facial expressions can _____ work in the opposite direction. According to this hypothesis, _____ from the facial muscles (feedback) are sent back to emotion _____ of the brain, and so a person's facial expression _____ influence

308

that person's emotional state. Consider Darwin's words: "The _____ expression by outward signs of an emotion intensifies it. _____ the other hand, the repression, as far as possible, _____ all outward signs softens our emotions." Can smiling give _____ to feelings of good will, for example, and frowning _____ anger?

Psychological research has given rise to some interesting _____ concerning the facial-feedback hypothesis. Causing participants in experiments _____ smile, for example, leads them to report more positive _____ and to rate cartoons (humorous drawings of people or _____) as being more humorous. When they are caused to _____, they rate cartoons as being more aggressive.

Appendix B: TOEFL Speaking Rubric

Score	General Description	Delivery	Language Use	Topic Development
4	The response fulfills the demands of the task, with at most minor lapses in completeness. It is highly intelligible and exhibits sustained, coherent discourse. A response at this level is characterized by all of the following:	Generally well-paced flow (fluid expression). Speech is clear. It may include minor lapses, or minor difficulties with pronunciation or intonation patterns, which do not affect overall intelligibility.	The response demonstrates effective use of grammar and vocabulary. It exhibits a fairly high degree of automaticity with good control of basic and complex structures (as appropriate). Some minor (or systematic) errors are noticeable but do not obscure meaning.	Response is sustained and sufficient to the task. It is generally well developed and coherent; relationships between ideas are clear (or clear progression of ideas).
3	The response addresses the task appropriately, but may fall short of being fully developed. It is generally intelligible and coherent, with some fluidity of expression though it exhibits some noticeable lapses in the expression of ideas. A response at this level is characterized by at least two of the following:	Speech is generally clear; with some fluidity of expression, though minor difficulties with pronunciation, intonation, or pacing are noticeable and may require listener effort at times (though overall intelligibility is not significantly affected).	The response demonstrates fairly automatic and effective use of grammar and vocabulary, and fairly coherent expression of relevant ideas. Response may exhibit some imprecise or inaccurate use of vocabulary or grammatical structures or be somewhat limited in the range of structures used. This may affect overall fluency, but it does not seriously interfere with the communication of the message.	Response is mostly coherent and sustained and conveys relevant ideas / information. Overall development is somewhat limited, usually lacks elaboration or specificity. Relationships between ideas may at times not be immediately clear.
2	The response addresses the task, but development of the topic is limited. It contains intelligible speech, although problems with delivery and/or overall coherence occur; meaning may be obscured in places. A response at this level is characterized by at least two of the following:	Speech is basically intelligible, though listener effort is needed because of unclear articulation, awkward intonation, or choppy rhythm / pace; meaning may be obscured in places.	The response demonstrates limited range and control of grammar and vocabulary. These limitations often prevent full expression of ideas. For the most part, only basic sentence structures are used successfully and spoken with fluidity. Structures and vocabulary may express mainly simple (short) and/or general propositions, with simple or unclear connections made among them (serial listing, conjunctions juxtaposition).	The response is connected to the task, though the number of ideas presented or the development of ideas is limited. Mostly basic ideas are expressed with limited elaboration (details and support). At times relevant substance may be vaguely expressed or repetitious. Connections of ideas may be unclear.
1	The response is very limited in content and/or coherence or is only minimally connected to the task, or speech is largely unintelligible. A response at this level is characterized by at least two of the following:	Consistent pronunciation, stress, and intonation difficulties cause considerable listener effort; delivery is choppy, fragmented, or telegraphic; frequent pauses and hesitations.	Range and control of grammar and vocabulary severely limit (or prevent) expression of ideas and connections among ideas. Some low-level responses may rely heavily on practiced or formulaic expressions.	Limited relevant content is expressed. The response generally lacks substance beyond expression of very basic ideas. Speaker may be unable to sustain speech to complete the task and may reply heavily on repetition of the prompt.
0	Speaker makes no attempt to respond OR response is unrelated to the topic.			

Source: http://www.ets.org/Media/Tests/TOEFL/pdf/Speaking_Rubrics.pdf

Appendix C: Sociolinguistic Interview Protocol

Age   Gender

1. Participant type

     ESL Student _____   Undergraduate _____   Graduate _____

     Career person _____   Other _____

2. When did the participant come to the United States?

     Less than 1 year 1-5 years   6-10 years   More than 10 years

3. Has the participant returned to Korea since then?

     Yes _____   No _____

4. If yes, how often does the participant return to Korea?

     Once a year ____   Twice a year ____   Every two years _____   Other _____

5. Why did the participant move to America?

     For graduate degrees _____   Career / job related _____

     To study English _____   Other _____

6. When did the participant first start learning English in Korea?

     Elementary _____   Middle school _____ Other _____

7. Did the participant receive private English lessons outside of school?

     Yes _____   No _____

8. If yes, when was it and for how long did the participant receive additional English lessons?

     Type: private tutoring _____   private English language schools _____

     Duration: 0-3 years _____   4-6 years _____   more than 7 years _____

9. Did the participant live in any English speaking country?

       Yes _____   No _____

10. If yes, when was it and for how long did the participant live there?

       _____

11. Did the participant have opportunities to interact with native English speakers in Korea?

       Yes _____   No _____

12. If yes, how often did the participant interact with native English speakers in Korea?

       Every day ____   A few hours a day ____   A few hours a week ____   Other ____

13. If enrolled in school, what kind of academic program or coursework is the participant doing?

       _____

14. If working, what kind of job does the participant have?

       _____

15. If enrolled in ESL program, what level is the participant placed at?

16. Where does the participant interact with native English speakers mostly?

       School _____   Work _____   Home _____   Other _____

17. How much time of the day does the participant use English?

       Less than 2 hours _____ 3-6 hours _____   More than 7 hours _____

18. What was the most difficult experience the participant had since living in the U.S.?

19. What was the most memorable trip for the participant in the U.S.?

20. Attitudes about English?

Appendix D: TOEFL Writing Rubric

Score	Task Description
5	**An essay at this level largely accomplishes all of the following:** • effectively addresses the topic and task • is well organized and well developed, using clearly appropriate explanations, exemplifications, and/or details • displays unity, progression, and coherence • displays consistent facility in the use of language, demonstrating syntactic variety, appropriate word choice, and idiomaticity, though it may have minor lexical or grammatical errors
4	**An essay at this level largely accomplishes all of the following:** • addresses the topic and task well, though some points may not be fully elaborated • is generally well organized and well developed, using appropriate and sufficient explanations, exemplifications, and/or details • displays unity, progression, and coherence, though it may contain occasional redundancy, digression, or unclear connections • displays consistent facility in the use of language, demonstrating syntactic variety and range of vocabulary, though it will probably have occasional noticeable minor errors in structure, word form, or use of idiomatic language that do not interfere with meaning
3	**An essay at this level is marked by one or more of the following:** • addresses the topic and task using somewhat developed explanations, exemplifications, and/or details • displays unity, progression, and coherence, though connection of ideas may be occasionally obscured • may demonstrate inconsistent facility in sentence formation and word choice that may result in lack of clarify and occasionally obscure meaning • may display accurate but limited range of syntactic structures and vocabulary
2	**An essay at this level may reveal one of more of the following weaknesses:** • limited development in response to the topic and task • inadequate organization or connection of ideas • inappropriate or insufficient exemplifications, explanations, or details to support or illustrate generalizations in response to the task • a noticeably inappropriate choice of words or word forms • an accumulation of errors in sentence structure and/or usage
1	**An essay at this level is seriously flawed by one or more of the following weaknesses:** • serious disorganization or underdevelopment • little or no detail, or irrelevant specifics, or questionable responsiveness to the task • serious and frequent errors in sentence structure or usage
0	**An essay at this level** merely copies words from the topic, rejects the topic, or is otherwise not connected to the topic, is written in a foreign language, consists of keystroke characters, or is blank.

Source: http://www.ets.org/Media/Tests/TOEFL/pdf/Writing_Rubrics.pdf

Appendix E: Writing Assessment Instrument

Direction: Please choose one of the following three questions and write an essay in response to the topic of your choice. There is a time limit of 30 minutes to write the essay.

- ▪ Do you agree or disagree with the following statement? Attending a live performance (for example, a play, concert, or sporting event) is more enjoyable than watching the same event on television. Use specific reasons and examples to support your opinion.

- ▪ Some people prefer to get up early in the morning and start the day's work. Others prefer to get up later in the day and work until late at night. Which do you prefer? Use specific reasons and examples to support your choice.

- ▪ Do you agree or disagree with the following statement? Children should begin learning a foreign language as soon as they start school. Use specific reasons and examples to support your position.

Appendix F: Information about the participants' background and language proficiency

Proficiency	Yrs. of stay in the U.S.	Age of arrival in the U.S.	CLOZE score	Speaking score	Writing score
Beginner					
1	0.4	27	2	2	1
2	0.6	22	13	1	1
3	1.2	24	15	2	2
4	0.3	23	19	1	2
5	3	24	19	1	1
6	1.5	23	22	2	3
7	0.7	23	23	2	2
8	0.8	24	23	2	2
9	0.2	21	31	1	1
10	0.8	26	32	1	2
Intermediate					
1	4.5	21	28	3	3
2	2	24	28	2	2
3	3	22	29	3	3
4	2.5	18	29	2	2
5	2	20	31	2	3
6	3	21	32	2	4
7	0.8	26	34	3	3
8	0.2	24	35	3	3
9	0.2	24	35	3	2
10	5	22	36	2	3
Advanced					
1	8	13	29	4	4
2	11	15	30	4	5
3	5.5	15	36	4	4
4	10	15	37	4	4
5	4	24	38	3	4
6	7	15	39	3	4
7	9	13	42	4	5
8	9	13	43	4	5
9	5	17	43	3	5
10	5	15	43	4	4

Appendix G: Script for the Oral Narrative Task

Thank you for taking time to sit down and talk with me.

First, I'd like to remind you that this session will be audio-recorded.

So how are you? How was your weekend? Did you have fun?

   [*Ice breaking conversation*]

Today, I'm going to ask you to tell me a story, or if you want, two or more stories: a story about some important event or incident in your life. You can think of an event, an incident or a day that was the most important or memorable to you in your life.

Can you tell me a story of some incident that left a great impact in your life?

   [*Give the participant enough time to think*]

(If no response) Well, has anything happened to you that left a big impact in your life?

(If still no response) How about your friends or family members? Can you tell me about something that happened to them?

(Finally) Then maybe you can think about an accident, your first boy friend / girl friend, or the most embarrassing moment in their life?

   [*If necessary, probing questions for more details, information and description will be asked.*]

Wow, that is a wonderful story. Thank you so much for sharing your story with me. It was great.

Appendix H: Script for the Written Narrative Task

You just told me a really great story. Thanks again for sharing them with me. Now, I'm going to ask you to write the stories with all the details as you just told me.

When you write, please assume that I have not heard the stories you just told me. I will forget your details, so please write as detailed as possible. Use this laptop computer to write your stories. If you want, you may write on paper. Dictionaries are not allowed. There is no time limit.

Direction: Please read the following dialogues carefully. For each <u>underlined</u> article in bold, determine whether each article is used appropriately. If you think the article use is appropriate, put a check mark next to OK. If you think it is not appropriate, write why and fix the article in the blank next to NOT OK.

(1)   *Phone conversation*

Grandmother: Oh, I just remember that John's turning two next week. What do

you think I should get for his birthday?

Mother: Well, these days, John really likes **the** toys that move. He likes to chase

after them.

OK _____   NOT OK _____

(2)   *In an office*

Alice: What did you do last night?

Robin: I went to a video store and got a German film and a video game. Then, I

came home and watched **the** film.

OK _____   NOT OK _____

(3)   *In a clothing store*

Clerk: May I help you?

Customer: Yes, please! I've rummaged through every stall, without any success. I

am looking for **a** warm hat. It's getting rather cold outside.

OK _____   NOT OK _____

(4)   *At a bookstore*

Chris: Well, I bought everything that I wanted. Are you ready to go?

Mike: Almost. Can you please wait a few minutes? I want to talk to **a** owner of

this bookstore. She is my old friend.

OK _____   NOT OK _____

(5)   *At a news office*

Reporter 1: Guess what? I finally got an important assignment.

Reporter 2: Great. What is it?

Reporter 1: This week, I'm interviewing **Ø** governor of Massachusetts, Mitt

Romney. I'm very excited!

OK _____   NOT OK _____

(6)   *At a family gathering*

Gary: I heard that you just started college. How do you like it?

Melissa: It's great. My classes are very interesting.

Gary: That's wonderful. And do you have fun outside of class?

Melissa: Yes. In fact, today, I'm having dinner with **the** girl from my class. Her

name is Angela and she is really nice.

OK _____   NOT OK _____

(7)   *At a neighbor store*

Mary: I heard that it was your son Roger's birthday last week. Did he have a good

celebration?

Roger: Yes, it was great. He got lots of gifts – books, toys. And best of all, he got

a puppy.

OK _____   NOT OK _____

(8)   *In an office*

Jane: After Thanksgiving week, I gained five pounds! Maybe I should sign up

for Weight Watchers or something.

Michelle: I know cutting down on calories will help, but **Ø** exercise on a regular

basis is the best way to lose weight and get healthy.

OK _____   NOT OK _____

(9)   *At a meeting on water quality report*

Town official: So did the test results come out? Is it safe to drink from the main

        reservoir in our town?

Water quality inspector: **0** water in the reservoir is polluted. I

        recommend that the city government bans residents from drinking it.

OK _____   NOT OK _____

(10)   *Phone conversation*

Mathilda: Hi, Sam. Is your roommate Lewis there?

Sam: No, he went to San Francisco for this weekend.

Mathilda: I see. I really need to talk to him. How can I reach him in San

        Francisco?

Sam: I don't know. He is staying with **the** mother of his best friend. I'm afraid I

        don't know who she is and I don't have her phone number.

OK _____   NOT OK _____

(11)   *At a gallery*

Sarah: Do you see that beautiful landscape painting?

Mary: Yes, it's wonderful.

Sarah: I would like to meet **an** artist of that painting. Unfortunately, I have no

        idea who it is, since the painting is not signed.

OK _____   NOT OK _____

(12)   *At home*

Karen: Where is Beth? Is she coming home for dinner?

Anne: No. She is eating dinner with **a** colleague. She didn't tell me who it is.

OK _____   NOT OK _____

(13)   *At a high school classroom*

Tom: Did you hear that George stood up against that bully for Mike yesterday?

Dave: I know! I was so impressed with **∅** courage that George took to help Mike.

    We were all very proud of him.

OK _____   NOT OK _____

(14)   *In a restaurant*

Waiter: Are you ready to order, sir? Or are you waiting for someone?

Client: Can you please come back in about twenty minutes? You see, I'm waiting.

    I'm planning to eat with **a** colleague from work. She will be here soon.

OK _____   NOT OK _____

(15)   *In an office*

Tom: It's not easy to come up with holiday gift ideas. It's especially hard for me

    to think about what to give to my little nephews and nieces. Do you have

    any advice for me?

Jane: In my case, I usually buy **∅** books for my nephews and nieces. I think

    reading is very important for children. I have not decided which books to

    buy this year, though.

OK _____   NOT OK _____

(16)   *At a college classroom*

Jane: I don't know why, but last night, I could not fall asleep.

Cindy: Mm.. did you eat anything for a late night snack?

Jane: Well, I had some hot chocolate and cookies.

Cindy: No wonder. You know that **the** chocolate contains caffeine, and I think

    that's why you could not fall asleep.

OK _____   NOT OK _____

Appendix J: The Text for the Fill-in-the-Article Test

Direction: Please read the following stories carefully and fill in the blanks with one of the following choices: (1) the (2) a (an) (3) ∅ . When you think that the blank cannot be filled with *a*, *an*, or *the*, then you need to insert ∅ . Your choice will be based on what you think is the most appropriate form for each blank in the context. While you make each article choice and write your choice in the blank, please think out loud and verbalize reasons for each choice. There is no time limit and you are allowed to think aloud in Korean, if you prefer to do so. This exercise will be audio-recorded.

Oscar hadn't eaten all day, but that was okay. Many scientific studies have concluded that (1)_____ less people eat, (2)_____ longer they live. These studies are based on (3)_____ experiments with (4)_____ mice and other small animals, so maybe (5)_____ conclusions don't apply to (6)_____ humans. But maybe they do. Oscar wanted to live to be 100. So, (7)_____ years ago, he had cut back from three meals (8)_____ day to two. Occasionally, he ate only one meal (9)_____ day. (10)_____ fewer meals Oscar ate, (11)_____ less guilty he felt about not exercising.

It was 11 p.m. Oscar peeled and ate (12)_____ banana. He filled (13)_____ pot with (14)_____ water, put (15)_____ pot on (16)_____ stovetop, and turned on (17)_____ gas. Then he sliced six radishes, salted them lightly, and ate them. Then he washed and ate two celery stalks. After that, he ate (18)_____ half cup of (19)_____ assorted nuts. He loved (20)_____ pecans best, although they were (21)_____ fewest in (22)_____ assortment. He sliced (23)_____ lemon in half, and then used (24)_____ squeezer to squeeze all the lemon juice into (25)_____ empty cup. When (26)_____ water in (27)_____ pot began to boil, he grabbed (28)_____ box of (29)_____ pasta from (30)_____ cabinet.

Oscar examined (31)_____ pasta box. On (32)_____ front, in (33)_____ large

322

letters, was "Barilla Penne—Italy's #1 Brand of Pasta." On (34)_____ side, in (35)_____ small print at (36)_____ very bottom, was "Product of U.S.A." Oscar shook his head; they tricked me again, he thought. He added two cups of (37)_____ pasta to (38)_____ boiling water and set (39)_____ timer. While (40)_____ pasta was cooking, Oscar took (41)_____ cucumber out of (42)_____ fridge. He peeled it, sliced it, salted (43)_____ slices, and ate them all while he read <u>This Week</u>, (44)_____ popular weekly news magazine. Finished with (45)_____ cucumber, he sliced (46)_____ pear and ate it.

(47)_____ timer went off. Oscar poured (48)_____ hot water through (49)_____ plastic strainer, and then put (50)_____ pasta into (51)_____ bowl. He added (52)_____ heaping tablespoon of (53)_____ butter to (54)_____ hot pasta and let it melt in. Then he poured (55)_____ lemon juice onto (56)_____ pasta. He chopped up part of (57)_____ red onion and mixed it into (58)_____ pasta. He drained (59)_____ can of (60)_____ boneless herring fillets and stirred them into (61)_____ pasta. He sprinkled (62)_____ garlic powder and grated (63)_____ parmesan cheese onto (64)_____ pasta. He put (65)_____ spoon into (66)_____ bowl, sat down at his dining room table, and dug in.

When Oscar finished (67)_____ pasta, he washed (68)_____ dirty dishes and utensils. Then he poured (69)_____ cold nonfat milk into (70)_____ large glass, sat down, and slowly ate half of (71)_____ five-ounce bar of Hershey's Special Dark candy. He washed it down with (72)_____ milk. Now, that was worth waiting for, he thought.

Appendix K: Think-Aloud Demonstration Script

Today, I will ask you to complete a simple exercise. I will give you a written story with blanks. Each blank can be filled by either (1) *a (an)* (3) *the*, and (4) ∅ . In other words, you have three choices for each blank. You need to fill each blank with one of the three choices. When you think that the blank does not require any particular form and it cannot be filled with *a* or *the*, then you need to insert ∅ .

While you read the text carefully and make a choice for each blank, I want you to explain to me why you are making such choices. I want you to be as specific as possible. You can take as much time as you want when you carefully consider each blank and think out loud so that I can understand exactly what goes through your mind.

To help you better understand how it will be done, I will play you a recording of someone who had to make choices for the following blanks. The task is about verb tenses and this is what this person did.

Sample

(1) It _____ more than two years to train a puppy to be a guide dog.

      (a) takes   (b) take   (c) took   (d) will take

(2) The first U.S. school for guide dogs _____ in 1929.

      (a) opens   (b) open   (c) opened (d) will open

(3) Now there _____ many schools for guide dogs.

      (a) is   (b) are   (c) was   (d) were   (e) will be

(4) Space shuttles _____ back and forth between Earth and space.

(a) goes   (b) go   (c) went   (d) will go

(5) The first one _____ on April 12, 1981.

(a) flies   (b) fly   (c) flyed   (d) flew   (e) will fly

(6) A space shuttle _____ four main parts.

(a) has   (b) have   (c) had   (d) will have

As you can see, she did her best to think out loud and express reasons behind each choice. *(Just like the way she did it in the demonstration recording, you are free to think out loud in Korean if it is more comfortable and easier for you to explain why you make certain choices.)* Please remember that I want to understand your thought process. I hope you will elaborate as much as possible your thought process and reasons behind each choice. There is no time limit, so please take as much time as you need to complete the task. This exercise will be audio-recorded.

Appendix L: Rates of Accuracy by Individual Participants

	Oral Narrative		Written Narrative		Fill-in-the-article test	
	the	*a*	*the*	*a*	*the*	*a*
Beginners						
1	0.0	7.1	11.8	0.0	56.8	30.0
2	85.7	36.4	100.0	50.0	91.9	25.0
3	21.4	16.7	53.8	27.3	35.1	80.0
4	69.2	30.4	71.4	28.6	81.1	85.0
5	50.0	41.7	45.5	65.0	65.6	75.0
6	69.2	25.0	66.7	83.3	70.3	50.0
7	38.5	43.8	27.8	13.3	81.1	90.0
8	60.0	25.0	55.6	7.7	70.3	95.0
9	64.9	14.3	66.7	56.3	83.8	95.0
10	78.8	17.6	100.0	40.9	67.6	90.0
Intermediates						
1	60.0	72.7	50.0	100.0	75.7	95.0
2	86.7	45.5	100.0	81.3	81.1	90.0
3	47.4	41.2	76.2	81.3	70.3	95.0
4	60.7	18.2	68.4	33.3	70.3	95.0
5	40.0	47.4	66.7	78.9	81.1	95.0
6	46.4	0.0	55.6	60.0	59.5	80.0
7	71.4	26.7	50.0	50.0	83.8	85.0
8	77.8	57.1	84.6	83.3	73.0	90.0
9	59.4	21.7	70.0	50.0	78.4	95.0
10	83.3	22.2	85.0	71.4	54.1	100.0
Advanced						
1	52.4	58.3	66.7	92.3	81.1	80.0
2	80.6	50.0	66.7	83.3	73.0	100.0
3	50.0	66.7	85.7	83.3	78.4	80.0
4	84.1	48.1	58.3	66.7	81.1	95.0
5	65.5	10.0	75.9	25.0	89.2	65.0
6	94.7	76.9	74.1	83.3	83.8	90.0
7	100.0	85.7	84.0	92.3	54.1	90.0
8	84.1	61.9	91.2	72.7	94.6	90.0
9	44.4	33.3	86.7	54.5	83.8	95.0
10	25.0	25.0	73.3	86.7	86.5	70.0

References

Akakura, M. A .(2009). *Effect of explicit instruction on implicit and explicit second language knowledge: An empirical study on English article acquisition*. Unpublished Ph.D. Dissertation, The University of Auckland.

Bohnacker, U. (1997). Determiner phrase and the debate on functional categories in early child language. *Language Acquisition, 6*(1), 49-90.

Bowles, M. A., & Leow, R. P. (2005). Reactivity and type of verbal report in SLA research methodology. *Studies in Second Language Acquisition, 27*, 415-440.

Brown, R. (1973). *A First Language: The Early Stages*. Cambridge, MA: Harvard University Press.

Butler, Y. G. (1999). *The role of metacognition in the development of the English article system among non-native speakers*. Unpublished Ph.D. Dissertation, Stanford University.

Butler, Y. G. (2002). Second language learners′ theories on the use of English articles. *Studies in Second Language Acquisition, 24*, 451-480.

Chesterman, A. (1991). *On definiteness: A study with special reference to English and Finnish*. Cambridge, UK: Cambridge University Press.

Christophersen, P. (1939). *The Articles: A Study of Their Theory and Use in English*. Copenhagen: Einar Munksgaard.

Cobbett, W. (1906). *Cobbett′s English Grammar*. London: Henry Frowde.

Dirven, R. (1989). Cognitive linguistics and pedagogical grammar. In G. Graustein & G. Leitner (Eds). *Reference Grammars and Modern Linguistic Theory*. Niemeyer: Tubigen, 56-75.

Ekiert, M. B. (2010). *Articles as expressions of definiteness in L2 English of Slavic speakers*. Unpublished Ph.D. Dissertation, Columbia University.

Ellis, R. (1991). Grammaticality judgments and second language acquisition. *Studies of Second Language Acquisition, 13*, 161-186.

Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford, UK: Oxford University Press.

Ellis, R. (2003). *Task-based Language Learning and Teaching*. Oxford, UK: Oxford University Press.

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning, 54*, 227-275.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition, 27*, 141-172.

Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effect of focused and unfocused written corrective feedback in an English as a foreign language context. *System, 36*, 353-371.

Ellis, R. & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition, 26*, 59-84.

Epstein, R. (2002). The definite article, accessibility, and the construction of discourse referents. *Cognitive Linguistics, 12*(4), 333-378.

Goad, H. & White, L. (2004). Ultimate attainment of L2 inflections: Effects of L1 prosodic structure, *EUROSAL Yearbook, 4*, 119-145.

Gor, K., & Long, M.H. (2009). Input and second language processing. In W.C.Ritchie & T.J.

Bhatia (eds.). Handbook of Second Language Acquisition. New York: Academic Press, 445-472.

Han, Z-H. (2008). On the role of meaning in focus on form. In Z-H. Han (Ed.), *Understanding second language process*. Clevedon, UK: Multilingual Matters, 45-79.

Han, Y., & Ellis, R. (1998). Implicit knowledge, explicit knowledge, and general language proficiency. *Language Teaching Research, 2*(1), 1-23.

Hawkins, J.A. (1978). *Definiteness and Indefiniteness*. London: Croom Helm.

Hawkins, J.A. (1991). On (in)definite articles: Implicatures and (un)grammaticality prediction. *Journal of Linguistics, 27*, 405-442.

Hawkins, R., Al-Eid, S., Almahboob, I., Athanasopoulos, P., Chaengchenki, R., Hu, J., Rezai, M., Jaensch, C., Jeon, Y., Jiang, A., Leung, Y.I., Matsunaga, K., Ortega, M., Sarko, G., Snape, N. & Velasco-Zarate, K. (2006). Accounting for English article interpretation by L2 speakers. *EUROSA Yearbook, 6*, 7-25.

Hewson, J. (1972). *Article and Noun in English*. The Hague: Mouton.

Huebner, T. (1979). Order-of-acquisition vs. dynamic paradigm: A comparison of method in interlanguage research. *TESOL Quarterly, 13*, 21-28.

Ionin, T., Ko, H. & Wexler, K. (2004). Article semantics in L2 acquisition: The role of specificity. *Language Acquisition, 12*(1), 3-69.

Ionin, T., Zubizarreta, M.L., & Maldonado, S.B. (2008). Sources of linguistic knowledge in the second language acquisition of English articles. *Lingua, 118*, 554-576.

Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic Influence in Language and Cognition*. New York: Routledge.

Jespersen, O. (1949). *A Modern English Grammar on Historical Principles*. London: Allen and Unwin.

Kim, G. S. & Ahn, G. S. (2003). Acquisition of the English definite article by Korean speakers. *English Language Teaching, 15*(4), 69-95.

Kim, N. K. (1990). Korean. In B. Comrie (Ed.). *The World′s Major Languages*. London, UK: Oxford University Press, 881-898.

Labov, W. (1984). Field methods of project on linguistic change and variation. In J. Baugh & J. Sherzer, (Eds.), *Language in use: Readings in sociolinguistics*. (pp. 28-53). Englewood Cliffs: Prentice Hall.

Labov, W. (1997). Some further steps in narrative analysis. *The Journal of Narrative and Life History, 7*, 395-415.

Langacker, R. W. (2008a). Cognitive Grammar as a basis for language instruction. In P. Robinson & N. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York: Routledge, 66-88.

Langacker, R. W. (2008b). *Cognitive Grammar, a Basic Introduction*. New York: Oxford University Press.

Lee, H. (1997). English article deletion in Korean EFL learners′ compositions. *Working Papers in Educational Linguistics, 13*(2), 41-50.

Liu, D., & Gleason, J.L. (2002). Acquisition of the article *the* by nonnative speakers of English: An analysis of four nongeneric uses. *Studies in Second Language Acquisition, 24*, 1-26.

Lyons, C. (1999). *Definiteness*. Cambridge, UK: Cambridge University Press.

Mandell, P. B. (1999). On the reliability of grammaticality judgment tests in second language

acquisition research. *Second Language Research, 15*(1), 73-99.

Master, P. (1990). Teaching the English articles as a binary system. *TESOL Quarterly, 24*(3), 461-478.

Master, P. (1997). The English article system: Acquisition, function, and pedagogy. *System, 25*, 215-232.

Master, P. (2002). Information structure and English article pedagogy. *System, 30*, 331-348

Meraji, S.R. (2011). Planning time, strategy use, and written task production in a pedagogic vs. a testing context. *Journal of Language Teaching and Research, 2*(2). 338-352.

Mizuno, M. (1999). Interlanguage analysis of the English article system: Some cognitive constraints facing the Japanese adult learners. *International Review of Applied Linguistics in Language Teaching, 37*(2), 127-153.

Oller, J.W. & Redding, E.Z. (1971). Article usage and other language skills. *Language Learning, 21*(1), 85-95.

Park, S. B. (2006). *The acquisition of written English articles by Korean learners*. Unpublished Ph.D. Dissertation, Southern Illinois University.

Parrish, B. (1987). A new look at methodologies in the study of article acquisition for learners of ESL. *Language Learning, 37*, 361-383.

Pica, T. (1983). The article in American English: what the textbooks don′t tell us. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and Language Acquisition* (pp. 222-233). London: Newbury House Publishers.

Prevost, P. & White, L. (2000).Missing surface inflection or impairment in second language acquisition?: Evidence from tense and agreement. *Second Language Research, 16*(2), 103-133.

Progovac, L. (1998). Determiner phrase in a language without determiners. *Journal of Linguistics, 34*, 165-179.

Reid, W. (1991). *Verb and Noun Number in English: A Fundamental Explanation*. New York: Longman.

Robertson, D. (2000). Variability in the use of the English article system by Chinese learners of English. *Second Language Research, 16*(2), 135-172.

Robinson, P. & Ellis, N. (2008). *Handbook of Cognitive Linguistics and Second Language Acquisition*. New York: Routledge.

Roehr, K. (2006). Metalinguistic knowledge in L2 task performance: A verbal protocol analysis. *Language Awareness, 15*(3), 180-198.
Roehr, K. (2007). Metalinguistic knowledge and language ability in university-level L2 learners. *Applied Linguistics, 29*, 173-190.

Schiffrin, D. (2006). *In Other Words*. Cambridge, UK: Cambridge University Press.

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Istanto, M. Nagami, J. W. Sew, T. Suthiwan, & I. Walker (Eds.), *Proceedings of CLaSIC 2010*, (pp. 721-737). Singapore: National University of Singapore, Center for Language Studies.

Skehan, P. (1996). A framework for the implementation of task based instruction. *Applied Linguistics, 17*, 38-62.

Sperber, D. & Wilson, D. (1986). *Relevance: Communication and Cognition*. Oxford: Blackwell.

Tarone, E. & Parrish, B. (1988). Task-related variation in interlanguage: The case of articles.

Language Learning, 38(1), 21-44.

Thomas, M. (1989). The acquisition of English articles by first- and second-language learners. *Applied Psycholinguistics, 10*, 335-355.

Trenkic, D. (2000). Establishing the definiteness status of referents in dialogue (in languages with and without articles). *Working Papers in English and Applied Linguistics, 7*, 107-131.

Trenkic, D. (2007). Variability in second language article production: Beyond the representational deficit vs. processing constraints debate. *Second Language Research, 23*(3), 289-327.

Trenkic, D. (2008). The representation of English articles in second language grammars: Determiners or adjectives? *Bilingualism: Language and Cognition*, *11*(1), 1-18.

Trenkic, D. (2009). Accounting for patterns of article omissions and substitutions in second language production. In M. Garcia Mayo & R. Hawkins (Eds.), *Second language acquisition of articles* (pp. 115-146). Amsterdam: John Benjamins.

White, B. (2009). Accounting for L2-English learners′ article choices. *MUS Working Papers in SLS, 1*(1), 14-37.

White, B. (2010). *In search of systematicity: A conceptual framework for the English article system*. Unpublished Ph.D. Dissertation, Michigan State University.

White, L. (2003). Fossilization in steady state L2 grammar: Persistent problems with inflectional morphology. *Bilingualism: Language and Cognition, 6*(2), 129-141.

Whitman, R.L. (1974). Teaching the article in English. *TESOL Quarterly, 8*(3), 253-262.

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics, 21*(4), 463-489.

Wray, A., & Perkins, M. (2000). The functions of formulaic language: An integrated model. *Language and Communication, 20*, 1-28.

Yoon, K.K. (1993). Challenging prototype description: Perceptions of noun countability and indefinite vs. zero. *International Review of Applied Linguistics in Language Teaching, 31*(4), 269-290.

Yotsukura, S. (1970). *The Articles in English: A Structural Analysis of Usage*. The Hague: Mouton.

Young, R. (1996). Form-function relations in articles in English interlanguage. In R. Bayley & D.R. Preston (Eds.). *Second Language Acquisition and Linguistics Variation*. Philadelphia, PA: John Benjamins B.V.

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.

Zegarac, V. (2004). Relevance theory and *the* in second language acquisition. *Second Language Research, 20*(3), 193-211.