การเรียนรู้การใช้คำนำหน้านามในภาษาอังกฤษของนักเรียนไทย โดยการวิเคราะห์ความรู้ด้านอภิภาษาศาสตร์

นางสาวสร้อยศิธร อิศรางกูร ณ อยุธยา

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาศิลปศาสตรดุษฎีบัณฑิต
สาขาวิชาภาษาอังกฤษเป็นภาษานานาชาติ (สหสาขาวิชา)
บัณฑิตวิทยาลัย จุฬาลงกรณ์มหาวิทยาลัย
ปีการศึกษา 2551

ACQUISITION OF THE ENGLISH ARTICLE SYSTEM BY THAI LEARNERS: AN ANALYSIS OF METALINGUISTIC KNOWLEDGE IN ENGLISH ARTICLE USE

Miss Soisithorn Isarankura

A Dissertation Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy Program in English as an International Language
(Interdisciplinary Program)
Graduate School
Chulalongkorn University
Academic Year 2008

Thesis Title	ACQUISITION OF THE ENGLISH ARTICLE SYSTEM BY THAI LEARNERS: AN ANALYSIS OF METALINGUISTIC KNOWLEDGE IN ENGLISH ARTICLE USE
By	Miss Soisithorn Isarankura
Field of Study	English as an International Language
Thesis Advisor	Assistant Professor Sudaporn Luksaneeyanawin, Ph.D.

Accepted by the Graduate School, Chulalongkorn University in Partial Fulfillment of the Requirements for the Doctoral Degree

..... Dean of the Graduate School
(Associate Professor Pornpote Piumsomboon, Ph.D.)

THESIS COMMITTEE

..... Chairman
(Assistant Professor Namtip Pingkarawat, Ph.D.)

..... Thesis Advisor
(Assistant Professor Sudaporn Luksaneeyanawin, Ph.D.)

..... External Member
(Assistant Professor Peansiri Vongvipanond, Ph.D.)

..... Member
(Assistant Professor Pavinee Thirakhupt, Ph.D.)

..... Member
(Pranee Modehiran, Ph.D.)

สร้อยศิธร อิศรางกูร ณ อยุธยา : การเรียนรู้การใช้คำนำหน้านามในภาษาอังกฤษของนักเรียนไทยโดยการวิเคราะห์ความรู้ด้านอภิภาษาศาสตร์. (ACQUISITION OF THE ENGLISH ARTICLE SYSTEM BY THAI LEARNERS: AN ANALYSIS OF METALINGUISTIC KNOWLEDGE IN ENGLISH ARTICLE USE) อ. ที่ปรึกษา: ผศ. ดร. สุดาพร ลักษณียนาวิน, 224 หน้า.

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษาความรู้ด้านอภิภาษาศาสตร์เกี่ยวกับระบบคำนำหน้านามในภาษาอังกฤษของนักเรียนไทย โดยวิเคราะห์ความรู้นี้จากหลักเกณฑ์และแนวทางที่ผู้เรียนใช้ในการเลือกคำนำหน้านามในบริบทต่างๆ ทั้งนี้เพื่อให้เข้าใจว่าเกณฑ์ที่ผู้เรียนนำมาใช้นั้น มีความเหมือนหรือแตกต่างกันอย่างไรกับเกณฑ์ที่เจ้าของภาษาใช้ และมีความสอดคล้องกับทฤษฎีที่นักวิจัยและนักภาษาศาสตร์กล่าวไว้เพียงใด การวิจัยนี้จะช่วยให้เข้าใจปัญหาของผู้เรียนในเชิงพัฒนาการตามแนวทางของทฤษฎีภาษาศาสตร์ กลุ่มตัวอย่างประกอบด้วยกลุ่มเจ้าของภาษา 1 กลุ่ม และกลุ่มนักเรียนไทย 2 กลุ่ม ซึ่งจำแนกโดยความสามารถทางภาษาอังกฤษเป็นกลุ่มที่มีความสามารถระดับสูงและกลุ่มที่มีความสามารถระดับต่ำ จำนวนกลุ่มละ 30 คน รวม 90 คน ผู้เข้าร่วมการวิจัยทั้งหมดทำแบบทดสอบโดยการเลือกเติมคำนำหน้านามลงในช่องว่างในบทความที่ใช้เป็นแบบทดสอบ ในทันทีที่ทำแบบทดสอบเสร็จกลุ่มตัวอย่างได้แสดงความรู้ด้านอภิภาษาศาสตร์โดยการให้เหตุผลของการเลือกใช้คำนำหน้านามนั้นๆโดยเขียนคำตอบลงในแบบสอบถาม การวิเคราะห์ข้อมูลในเชิงคุณภาพอาศัยเกณฑ์การจำแนกนามวลีตามทฤษฎีของฮิว์บเนอร์ (Huebner, 1983) ส่วนการวิเคราะห์ข้อมูลเชิงปริมาณใช้ทั้งสถิติเชิงพรรณนาและสถิติเชิงอนุมาน

ผลการวิจัยพบว่า นักเรียนกลุ่มที่มีความสามารถทางภาษาสูงส่วนใหญ่จะใช้ความรู้ด้านอภิภาษาศาสตร์เชิงวัจนปฏิบัติในการพิจารณาคำนำหน้านามซึ่งสอดคล้องกับกลุ่มเจ้าของภาษา แต่จากการที่นักเรียนในกลุ่มนี้มักจะนำข้อสมมติฐานของตนเข้ามาประกอบในการพิจารณามากเกินไป จึงทำให้อนุมานว่าสิ่งที่ผู้พูดหรือผู้เขียนกล่าวถึงน่าจะเป็นที่เข้าใจกันในกลุ่มของผู้ฟังหรือผู้อ่าน ดังนั้นจึงเลือกใช้คำนำหน้านาม 'the' มากเกินความจำเป็น ส่วนนักเรียนกลุ่มที่มีความสามารถทางภาษาต่ำ มักอาศัยความรู้เกี่ยวกับโครงสร้างข้างเคียงในประโยคมาใช้เป็นหลักในการเลือกคำนำหน้านาม เนื่องจากการใช้ 'the' ต้องพิจารณาโดยอาศัยบริบทเป็นหลัก นักเรียนกลุ่มความสามารถทางภาษาต่ำจึงมักเลือกใช้ 'a' ในทุกบริบทที่มีความเป็นไปได้ ในการให้เหตุผลของนักเรียนทั้งสองกลุ่มพบว่าสมมติฐานมีความหลากหลาย บางสมมติฐานก็ตรงกับที่เจ้าของภาษาใช้ บางสมมติฐานก็แตกต่าง ทั้งนี้อาจสืบเนื่องมาจากประสบการณ์การเรียนภาษาอังกฤษที่แตกต่างกัน ซึ่งทำให้การเลือกใช้คำนำหน้านามมีความแตกต่างระหว่างกลุ่มอย่างมีนัยสำคัญทางสถิติ นอกจากนี้ผลการวิจัยยังสรุปได้ว่านักเรียนทั้งสองกลุ่มมีปัญหากับบริบทที่ไม่ต้องใช้คำนำหน้านาม (หรือการใช้ 'Ø') กล่าวโดยรวมนักเรียนใช้ 'Ø' ในอัตราส่วนที่น้อยมาก อีกทั้งนักเรียนไม่สามารถระบุเหตุผลที่ชัดเจนมาอธิบายการใช้ 'Ø' ได้ การวิจัยนี้ได้ชี้ให้เห็นถึงสาเหตุของปัญหาด้านการใช้คำนำหน้านามของนักเรียนไทยที่มีความสามารถทางภาษาต่างกัน ซึ่งจะเป็นแนวทางในการพัฒนาวิธีสอนให้กับนักเรียนไทยได้อย่างมีประสิทธิภาพมากยิ่งขึ้น

สาขาวิชา: ภาษาอังกฤษเป็นภาษานานาชาติ ลายมือชื่อนิสิต.....

ปีการศึกษา: 2551 ลายมือชื่ออาจารย์ที่ปรึกษา.....

4889679820 : MAJOR ENGLISH AS AN INTERNATIONAL LANGUAGE
KEY WORD: METALINGUISTIC KNOWLEDGE / ENGLISH ARTICLE SYSTEM / SECOND LANGUAGE ACQUISITION (SLA) / INTERLANGUAGE

SOISITHORN ISARANKURA : ACQUISITION OF THE ENGLISH ARTICLE SYSTEM BY THAI LEARNERS: AN ANALYSIS OF METALINGUISTIC KNOWLEDGE IN ENGLISH ARTICLE USE. THESIS ADVISOR : ASST. PROF. SUDAPORN LUKSANEEYANAWIN, PH.D., 224 pp.

This study investigated metalinguistic knowledge employed by Thai learners when making article choices in given contexts. By examining what rules were most commonly formulated and used in determining English articles, the study aimed to understand whether there were significant discrepancies between the rules and/or explanations upon which learners based their judgments and those that were consistent with linguistic theory. Moreover, in order to better understand the nature of learners' problems at different stages in the evolving interlanguage, the study explored the differences in metalinguistic knowledge between native English speakers and Thai learners at high and low proficiency levels. Thirty native English speakers and sixty Thai students, divided into high and low proficiency groups, participated in the study. Immediately following a fill-in-the-article test, the participants offered reasons for their article choices on a questionnaire. The data were analyzed quantitatively, using descriptive and inferential statistics. Qualitative analyses were also performed across noun phrase classifications based on Huebner's (1983) model.

The results indicate that the metalinguistic knowledge of the high English proficiency group was primarily pragmatically-oriented and more consistent with that of the native English speaker participants. However, as they incorporated too many assumptions to their readings, the high-proficiency students often wrongly judged the references to be part of the assumed hearer's knowledge, and thus they overused the definite article 'the'. The students with low proficiency relied heavily on limited structural cues in choosing articles. As 'the' was more deeply context-dependent, the low-proficiency students often used 'a' wherever possible. It was found that the hypotheses formulated by the Thai students of both proficiency groups varied probably due to the type and amount of input (i.e. language instruction) that the students had received. Some hypotheses were consistent with those of native speakers, while others were not. This resulted in the statistically significant differences in performance on the article test between the two proficiency groups. The 'zero' article was found to be most problematic for both groups of Thai students and was used the least. Furthermore, the Thai students had difficulties articulating a reason to support their choice of the 'zero' article. The findings reveal where article problems in Thai students with different proficiency levels originate. Based on the findings of the study, pedagogical implications are given so that a more effective model could be developed for teaching the English article system to Thai students.

Field of study: English as an International Language Student's signature.....

Academic year: 2008 Advisor's signature.....

ACKNOWLEDGEMENTS

I wish to express my gratitude to a great number of people who guided me and helped me conduct this study. Without their support, I wouldn't have been able to bring this dissertation to its completion. First, my profound appreciation goes to my committee members: Assistant Professor Dr. Namtip Pingkarawat, Assistant Professor Dr. Pavinee Thirakhupt, Dr. Pranee Modehiran, and Assistant Professor Dr. Peansiri Vongvipanond.

I am particularly indebted to my advisor, Assistant Professor Dr. Sudaporn Luksaneeyanawin, for her insightful comments and suggestions, and especially her endless patience and unfailing encouragement, which helped push me through all the difficult stages of this dissertation.

Special thanks go to Assistant Professor Dr. Cindy M. McClenagan for her kind assistance in allowing me to collect data from her students in two of her classes at Wayland Baptist University. I also wish to thank the participants for their cooperation and willingness to offer all requested information during and after the test administration. I owe a special word of thanks to Eddie and Sue Turner and Kriwin Sarawichitr, who gave all the possible support when I collected data in the United States.

My sincere thanks also go to several people at San Jose State University: Professor Emeritus Dr. Peter Master, Professor Dr. Thom Huebner, Associate Professor Dr. Swathi Vanniarajan, Karen O'Neill and Marianne Wheeler, for their supportive and encouraging advice during my three-month stay in the United States.

Finally, I am deeply grateful to all professors at EIL who provided me with invaluable theoretical and practical knowledge. I was very fortunate to benefit greatly from their wisdom and expertise throughout my entire studies in the program. I also wish to acknowledge the productive contributions of the faculty members and friends for sharing their experiences and suggestions, and for making my years at EIL the most enjoyable and unforgettable part of my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I
INTRODUCTION

1.1 Background of the Study

The English article system, which includes *a/an*, *the,* and the *zero* article *Ø*, is a notoriously complex aspect of English grammar. Although article usage entails the use of seemingly simple morphemes, teaching and learning how to employ these morphemes is not as simple as it might appear. Many advanced nonnative learners of English have difficulties in using articles properly even when all other structural elements of the language have been mastered. In the pedagogical literature, teachers of English as a second or foreign language (ESL/EFL) report that articles are often their number one difficulty (Celce-Murcia and Larsen-Freeman, 1999). Teaching the English article system is problematic for both nonnative and native English teachers. Many non-native English teachers admit that they do not fully understand how to use English articles correctly themselves. If this is the case, one cannot expect them to be able to teach their students how to do so. For native English teachers, teaching English articles to ESL/EFL learners is also not an easy task. It may be true that native English speakers can detect errors in their students' article usage without much difficulty. However, given that native speakers make article choices almost intuitively, they often find it difficult to explain to their students why this article is appropriate and why others are not.

Researchers have attempted to analyze the difficulties in article acquisition and investigate various methods of teaching the article system to speakers of other languages. Yet, there appears to be no satisfactory or convincing way of dealing with these difficulties. Why is the English article system so difficult for second language (L2) learners to acquire? Researchers and teachers ascribe the difficulty in article acquisition to numerous factors. One of the most widely accepted explanations would be the complexity of the English article system itself (Young, 1996; Master, 2002). The strikingly complex usage of English articles stems from multiple form-function mappings. It thus requires learners to go beyond the one-form-one-function strategy, which is a commonly practiced method of learning among less advanced language learners. This explains why acquisition is usually delayed until a late stage of L2 development.

A second possible explanation can be attributed to the fact that, like most function words, articles are generally overlooked by learners and teachers when processing language primarily for meaning (Ekiert, 2004). The articles rarely cause misunderstanding when misused in spoken language (Master, 1990). Accordingly, they are not usually regarded as crucial communicative devices. In fact, articles can be very influential on communication. The misuse of articles can lead the listener to a different interpretation than intended and can be the cause of serious miscommunication.

Although there has been a considerable amount of research that has investigated L2 article acquisition, the findings reveal contradictory results in acquisition patterns among learners of various first/native language (L1) backgrounds. Many findings (e.g., Huebner, 1983; Master, 1997; Parrish, 1987; Thomas, 1989) suggest the integration of the definite article *the* into the learner's interlanguage before the integration of the indefinite article *a*. Others (e.g. Liu & Gleason (2002), Young (1996), and Ekiert (2004) report opposing results: that the accurate control of *a* occurred before the mastery of *the*. Brown (1983) attributes the inconsistency to the variety of methodological approaches used across studies (e.g. longitudinal, pseudolongitudinal, and cross-sectional approaches in data-gathering procedures).

Another factor responsible for inconsistency in learner accuracy concerns the fact that the participants represent a multitude of various L1 speakers: Chinese, Japanese, Korean, Finnish, Polish, German, Spanish, and Russian, among others. L2 learners whose L1s contain an article system [+ART] appear to differ in English article acquisition from those whose L1s lack such a system [-ART]. Differences exist not only in their learning conditions and learning strategies, but also in their well-established knowledge about how languages work. Thus they often utilize or transfer much of the knowledge of their source language (L1) to help map certain grammatical features of the target language (L2). Master (1988) maintains that even when learners' L1s have articles, the differences from English in the application of the rules and article usage can also cause difficulty for L2 learners. Problems are even exacerbated when learners' native languages lack an article system. Larsen-Freeman (1975) reports in her study that the rank of articles in morpheme acquisition order was lower among Japanese learners than for those of other L1 backgrounds. Japanese does not have independent articles or article-like morphemes. Similarly, the Thai language does not have functional

equivalents of the English articles. Correspondingly, there is a great deal of observational evidence that Thai learners have difficulty with the English article system (e.g. Iamchote, 1971; Supol, 1980; Srioutai, 2001). In particular, these learners often omit the article where native English speakers would use one. Such errors are viewed to be influenced by L1 transfer.

Added to these factors, researchers' varied interpretations in learners' article choice can be attributed to the inconsistency of accuracy across studies. Learners' overproduction of the *zero* article (*Ø*) presents a clear example. While Parrish (1987) and Master (1997) interpret *Ø* overuse as an indication of mastery of the *zero* article in initial stages of L2 acquisition, Thomas (1989) reports that the phenomenon in which her [-ART] participants "produced the *zero* article more frequently in *a* and *the* contexts" was rather "failure to use any article" (p. 349). Thomas concludes that *Ø* overuse, or equivalently, failure to use articles, is a reflection of the differences between the target language and the native language, or any other language previously acquired. She views the delayed acquisition of *Ø* to be the result of L1 transfer.

Differences in English article acquisition among L1 children and L2 learners are indeed intriguing. There is strong empirical support that L1 children acquire the article system at an early age, typically between the age of three and four (Emslie & Stevenson, 1981; Garton, 1983; Maratsos, 1976, as cited in Cziko, 1986). One possible explanation lies in the differences in learning environments and availability of input between L1 children and L2 learners. Although it is still unclear as to how the acquisition mechanism of the article system works among L1 children, one may assume that it is somehow different from the acquisition mechanism employed by L2 learners. While native speakers' article choices, in most cases, appear to be an unconscious process, most L2 learners require some kind of conscious mechanism in generalizing or retrieving rules given by a teacher under supervised conditions.

This study aimed to investigate the kinds of rules Thai learners employ in determining an article in a given situation; how such rules are formulated; and how consistently these rules are applied in certain contexts as opposed to the mental representations used by native English speakers in choosing articles. It is claimed that native speakers often make article choices based on their intuitive evaluation of the

context in question and find it hard to articulate rules or explanations for their article usage. However, one may assume that native speakers who have received formal instruction regarding article use in school can articulate some rules or reasons for their article choices. By comparing mental representations (or "rules") of the English article system between native English speakers and Thai learners of varying proficiency levels, I believe that teachers will be able to better understand learners' problems with article usage. It seems that without a well-defined means of identifying and understanding learners' problems, teaching the article system effectively to Thai learners remains an elusive goal. This study thus offers an alternative approach to addressing this issue by examining metalinguistic (explicit) knowledge employed by native English speakers and Thai learners at high and low proficiency levels in deciding which article to use in various noun phrase (NP) contexts. It is hoped that the findings of the study will provide some evidence regarding the nature of EFL learners' problems, areas of difficulty, and the theories learners actively employ in the management of their article selection. Moreover, learners' interpretations of the English article system are expected to reveal their article acquisition processes at different stages in their interlanguage development. The information gained from this study is hoped to provide a basis for developing a more effective model for teaching the English article system.

The organization of this dissertation is as follows. In this chapter, research questions, objectives, hypotheses, and key term definitions are presented. In Chapter 2, I review the literature relevant to the study. Chapter 3 provides the linguistic background of Huebner's (1983) NP classification model, which is the framework used in the study. Chapter 4 discusses research methodology. In Chapter 5 and Chapter 6, I present the quantitative and qualitative results of this study. Chapter 7 provides the conclusions and some implications for future EFL instruction of the English article system.

1.2 Research Questions

The goal of this study was to examine metalinguistic knowledge employed by Thai learners, as opposed to native English speakers, in understanding the English article system. In so doing, conceptual differences with regard to pragmatic, semantic, and syntactic considerations for article usage could be identified. Therefore, this study aimed to investigate the following questions:

(1) What is the metalinguistic knowledge used by native English speakers and Thai EFL learners of high and low proficiency levels with regard to the English article system?

(2) What are the similarities and differences in the metalinguistic knowledge formulated and used in determining English articles by native English speakers and Thai learners of high and low proficiency levels?

(3) What are the problems underlying the use of English articles for Thai learners with regard to the metalinguistic knowledge of definiteness, genericity, and countability?

1.3 Objectives of the Study

The objectives of the study were:

(1) To examine the metalinguistic knowledge of the English article system in native English speakers and Thai EFL learners of high and low proficiency levels;

(2) To compare and contrast metalinguistic knowledge formulated and used to determine English articles between native English speakers and Thai learners of high and low proficiency levels;

(3) To identify the problems of Thai learners of high and low proficiency levels in using English articles.

1.4 Statement of Hypotheses

To carry out the objectives of this study, the following hypotheses were formulated and tested:

Hypothesis 1:

Native English speakers' metalinguistic knowledge of the English article system is relatively stable and pragmatically-oriented. Thai learners' metalinguistic knowledge is primarily syntactically and semantically-oriented. However, the metalinguistic knowledge of Thai learners of high proficiency is more pragmatically-oriented than learners of low proficiency levels.

Hypothesis 2:

The metalinguistic knowledge of the English article system formulated and used by native English speakers and Thai learners is similar syntactically and semantically, but different pragmatically.

Hypothesis 3:

The problems underlying the use of English articles for Thai learners are mostly pragmatically-oriented. Problems also arise from the surface syntactic structures of noun phrases in English.

1.5 Scope of the Study

Native English speakers' metalinguistic knowledge explored in the study was used to compare and contrast with Thai learners' metalinguistic knowledge for determining article use. The research was undertaken as follows.

(1) In addition to the attempt to identify Thai learners' potential problems in using English articles, this study aimed to examine the developmental pattern of the article acquisition processes at different stages of their interlanguage. Thus, data-gathering procedures were conducted in a cross-sectional design using two groups of Thai learners at high and low proficiency levels and one group of native English speakers. Research methodology will be discussed in more detail in Chapter 4.

(2) To eliminate the influence of familiarity with linguistic issues on the metalinguistic reasoning, native English speakers recruited for the study were undergraduate American students majoring in fields other than English, language studies, and linguistics. All the Thai learners had received formal English instruction in EFL settings in Thailand, and none had resided in an English speaking country for longer than three months.

(3) The framework in the study is based mainly on Huebner's (1983) classification model, which itself is based on Bickerton's (1981) semantic and discourse universal framework. The model will be discussed in more depth in Chapter 3.

(4) As proper nouns, idioms and formulaic expressions are often learned as chunks with or without a particular article, these NPs were not included in the analysis. Furthermore, the use of articles in these contexts seems unlikely to fall under the same rules as for other NPs.

1.6 Limitations of the Study

Although the present study aimed to provide a clearer picture of the ways in which Thai learners acquire the English article system, it had the following limitations.

(1) The study focused on article acquisition and difficulties of Thai learners with relative homogeneity in their EFL learning environment. Therefore, generalizability of the findings of this study may be limited, and not completely feasible to other populations.

(2) This study's cross-sectional design may be viewed as reflecting only one particular phase of the interlanguage development, which limits one's ability to make strong inferences about learners' sequential process in article usage across proficiency levels in their evolving interlanguage. However, the metalinguistic reasoning rendered by participants provides confirmatory data of language performance in relation to rules or theories employed by learners across groups. The findings thus present useful insights into learners' article acquisition order to a large extent.

(3) The third limitation concerns the accessibility of metalinguistic knowledge for article use when participants perform a given task. Krashen (1976) argues that 'hard' rules are 'acquired' and thus may be represented in implicit knowledge, whereas 'easy' rules are 'learned' and may be represented in explicit knowledge. As the English article system is considered "hard" grammar, it is not clear to what extent participants actually access the metalinguistic knowledge and how well they can articulate such knowledge. Although participants in the present study articulated metalinguistic reasoning quite well, this does not mean that it fully reflected their implicit knowledge of the English article system. The role of implicit knowledge in relation to explicit knowledge is still an open question. This issue will be discussed further in the following chapter.

1.7 Definition of Terms

For the purpose of the study, the following terms are thus defined.

(1) Second language acquisition (SLA)

The term 'second language acquisition', referred to in many studies as 'SLA' or 'L2 acquisition', is used in this study to refer to acquisition of an additional language after one has acquired his/her first or native language. It is not meant to contrast with foreign language (FL) acquisition. L2 acquisition used in this study refers to Thai learners' acquisition of the English language.

English as a second language (ESL) and English as a foreign language (EFL)

The purpose of contrasting ESL with EFL is based on the open question of whether the English article system is teachable and learnable through formal English instruction or if it is unteachable and can only be acquired through exposure.

Thus, in this study, 'English as a Second Language' (ESL) is defined as English language learning that takes place in a country where English is spoken natively. For example, a speaker of Thai learning English in countries such as the USA, the UK, Australia, or New Zealand would be categorized as an ESL learner.

'English as a Foreign Language' refers to English language learning that takes place where English is not the native language of the society, and where learners have few opportunities to practice the target language outside the classroom. This is a common situation in countries such as Thailand, Korea, or Japan, where learning English is usually confined to the classroom.

However, distinguishing between EFL and ESL is not always simple. The formerly clear distinction between the two contexts has been made less discernible by the current trend toward globalized communities. The penetration of English-based media (e.g. television, the Internet) provides ready access to English even in somewhat isolated settings. This has led to English being regarded not as a second language or foreign language, but as an international language (EIL). Despite the fact that EFL is becoming quite complicated to define, I still would like to classify Thai learners of English as EFL

learners based on the hypothesis that the way in which acquisition proceeds in EFL and ESL situations might be different. The English article system is a linguistic form that has a high likelihood to be influenced by learning environment differences. The main aim of identifying EFL contexts for Thai learners in this study is, therefore, to establish a clear distinction in the metalinguistic knowledge of native English speakers and Thai learners who have learned English through formal instruction in EFL settings in Thailand.

(3) Metalinguistic knowledge

Metalinguistic knowledge concerns the mechanism underlying the performance of language activities. It can be generally defined as learners' explicit knowledge about language (e.g. Bialystok, 1979). Explicit knowledge is conscious and declarative knowledge that can be brought into awareness and that is potentially available for verbal report (Hulstijn, 2005). Metalinguistic knowledge includes statements about intuitions of grammaticality, opinions, perceptions of the utterance and the abstract knowledge about the language, its structure and its uses.

In this study, metalinguistic knowledge is used to refer to the awareness and the ability to think or verbalize explicitly about the usage of English articles in terms of their syntactic, semantic, and pragmatic features underlying article usage in different contexts. The metalinguistic tasks demand an explanation of applicable rules for each article use in contextualized texts. The language can consist of technical or semi-technical terminology.

Metalinguistic theory and research will be discussed in more detail in Chapter 2.

1.8 Significance of the Study

The present research is worth conducting on several counts.

(1) There has been no study to date that investigates metalinguistic knowledge in English article usage employed by Thai learners as opposed to native English speakers. This study aims to offer one approach to exploring learners' problems in article acquisition. It is hoped that analyses of metalinguistic knowledge will give better insights

into a number of syntactic, semantic, pragmatic and conceptual differences between native English speakers and Thai learners with regard to article use in a given context.

(2) This study seeks to provide evidence that pragmatic orientations play a vital role in native English speakers' metalinguistic knowledge of the English article system. Basic rules of English articles provided in grammar books are mainly syntactically and semantically-based. Pragmatic functions are not extensively described in grammar textbooks. I believe that it is vital for L2 learners to have metalinguistic awareness and knowledge that is consistent with the native speaker's competence in order to be able to master the English article system.

(3) It has been suggested that learners with different proficiencies have different kinds of problems with articles. Yet, there has been no investigation in any great detail of the different theories in article use that learners may formulate at different stages of their evolving interlanguage. There are some plausible psycholinguistic explanations of learners' misuse or overuse of articles—e.g. interference from the learners' L1 (Master, 1997; Thomas, 1989). However, there is no guarantee that they are in fact the only sources of the non-targetlike uses observed in L2 learners. In order for teachers to correct errors that learners make in their article use, it is extremely important to understand how learners actually perceive the article system and internalize it.

The pedagogical implications for students with different proficiency levels should accordingly be clear. Additional sources of information such as learners' metalinguistic knowledge of English article use can be a valuable tool to determine how learners of varying proficiency levels comprehend the article system and what specific problems are present at different stages. This will allow EFL teachers to develop a more informed model of teaching the English article system and thus design more appropriate instructional materials especially for Thai learners.

CHAPTER II
REVIEW OF LITERATURE

This chapter reviews theories and research studies that are relevant to the present study. It consists of five sections. In the first section, I describe linguistic aspects underlying English articles: namely, definiteness, generic/specific reference, and countability. Section 2.2 discusses these three aspects in Thai nouns/noun phrases. In Sections 2.3, I review studies on first language (L1) article acquisition. This is followed by the literature review of second language (L2) article acquisition. In the final section, theories and research on metalinguistics are presented.

2.1 Linguistic Aspects of English Articles

Researchers and teachers have attempted to identify the causes of article difficulty for L2 learners and suggested that the difficulty in article usage is primarily caused by the complexity of the English article system itself. The complexity in article usage stems from the multiple concepts underlying English noun phrases—namely, (1) what sense the speaker has of the hearer's knowledge about the topic (definite vs. indefinite); (2) whether the speaker is referring to an entity in a specific place and time, or generically to all such entities at any place or time (specific vs. generic reference); and (3) whether a noun is a discrete item or a mass (count vs. mass). Past research shows that these concepts attribute to the difficulties in L2 article acquisition to a certain extent.

This section discusses semantic and pragmatic functions which underlie English noun phrases and determine the choices of articles. Three main concepts are presented: (1) definiteness (2) generic reference and (3) countability.

2.1.1 Definiteness

2.1.1.1 Definite Descriptions

Definite descriptions are one of the most common constructs in English and have been extensively studied by linguists, psychologists, and computational linguists (e.g. Russell, 1905; Christophersen, 1939, Hawkins, 1978, Poesio & Vieira, 1998). Master (2002) has described the distinction between definite and indefinite noun phrases (NPs)

in terms of 'the sense the speaker/writer has of the hearer/reader's knowledge about the topic'. In assessing a new listener's knowledge, the speaker's decision of whether to use the definite or indefinite article depends on subtle, perhaps intuitive 'readings' of what the listener knows (ibid). Kearns (2000) proposes that definite noun phrases pick out particular objects that the hearer can identify. Generally, discussion of definite descriptions concentrates on noun phrases beginning with the definite article *the*[1]. According to Quirk et al. (1985), who incorporate the insights of Hawkins (1978), noun phrases with *the* refer to something which can be identified uniquely in the contextual or general knowledge shared by speaker and hearer.

Lyons (1980) claims that articles have been recognized by linguists to be a fundamental element that links the sentence to the situation of communication. According to Lyons, when communication is concerned, one has to take into account pragmatic considerations in order to use English articles correctly. The definite article, in particular, is treated under most pragmatic accounts in terms of its role in reference. There are several terms used by linguists to define the property of definite descriptions; for example, *uniqueness* (Russell 1905; Neale 1990), *familiarity* (Christophersen 1939; Hawkins 1978; Prince 1981), *inclusiveness* (Hawkins, 1991), and *location* (Hawkins, 1978; Lyons, 1980). Among these theories, the work of Hawkins (1978) on definite description uses has been referred to in many research studies. Hawkins' account is more toward a pragmatic approach. He describes the relationship between semantics and pragmatics, and more generally human reasoning in everyday language use, and proposes that articles carry conversational implicatures that are responsible for the interpretations that they exhibit in the use of the language. According to Hawkins (1978), the use of the definite article involves the performance of three speech acts—i.e. the speaker (a) introduces a referent[2] to the hearer; (b) instructs the hearer to locate the referent in some shared set of objects; and (c) refers to the totality of the objects or mass within this set which satisfy the referring expression (cited in Lyons, 1980: 83).

---

[1] Concentration will be only on NPs with the definite article *the*, not with other cases of definite noun phrases such as pronouns, or possessive descriptions.

[2] The term *referent* is used to indicate the object in the world that is contributed to the meaning of an utterance by a definite description—e.g., when people say *George W. Bush*, it is the referent of a referential use of the definite description *the president of the USA between 2001 and 2008*.

Hawkins (1978) further develops and extends Christophersen's (1939) list and identifies the following classes or uses of definite descriptions:

Anaphoric Use. These are definite descriptions that co-specify with a discourse entity already introduced in the discourse or its antecedent in a text. The definite description may use the same head as its antecedent, or any other words capable of indicating the same antecedent (e.g. a synonym, a hyponym, etc.).

(1) a. Bill was working at *a lathe*…Suddenly *the lathe* stopped.
b. Bill was working at *a lathe*…Suddenly *the machine* stopped.
c. Fred *traveled to Munich*…*The journey* was long and tiring.
(Lyons, 1980:84)

In (1)a the anaphoric NP *the lathe* refers to the previously mentioned NP *a lathe* in the discourse, while the anaphoric NP *the machine* in (1)b is not identical to the antecedent, and the NP *the journey* in (1)c refers to the previously mentioned verb phrase *traveled to Munich*.

Immediate Situation Uses. These are definite descriptions which refer to an object in the situation of utterance. The referent may be visible, or its presence may be known to (or inferred by) both the speaker and hearer. The *visible situation use* occurs when the object referred to is in the field of vision to both the speaker and hearer, as in:

(2) a. Pass me *the bucket*.
b. Close *the door*.
c. Put out *the light*.

In (2)a, *Pass me the bucket* will be unambiguous for the hearer if there is just one bucket visible at the time of utterance. However, in (2)b and (2)c, although there may be several doors or several lights, the sentences can still be unambiguous when the context allows for the hearer to understand which one is intended by the speaker.

Definite descriptions can also be used when a referent is a constituent of the immediate situation in which the use of the definite description is located, without necessarily being visible. Hawkins classifies those definite descriptions as *immediate situation uses*, as in:

(3) a. Don't feed *the pony*.
b. Beware of *the dog*. (Lyons, 1980:84)

As a sign on a gate, *the* in (3)a and (3)b informs the reader of the existence of a pony and a dog, and perhaps instructs him/her to use the situation to find it.

Larger Situation Uses. A larger situation set of entities is characterized by the fact that even though a NP represents a first-mention of some object, its reference can be uniquely identified based on the shared knowledge between the speaker and hearer, which does not derive from the immediate situation. Hawkins lists two uses of definite descriptions which serve as shared knowledge among members of the same community in the non-immediate or larger situation of utterance. The first use of definite descriptions may rely on *specific knowledge about the larger situation*. Thus, the shared knowledge of inhabitants of the same village enables them to talk about *the pub,* or *the church*; inhabitants of the same town speak of *the mayor*; and inhabitants of the same country can do likewise with *the prime minister, the president,* or *the king* and *the queen*, as in:

(4) Mrs. Dobson has left for *the church*.
(5) *The prime minister* will give an opening speech.

Specific knowledge is not always a necessary part of the meaning of larger situation uses of definite descriptions. General knowledge about the existence of certain types of objects in certain types of situations may be sufficient. Hawkins classifies the second use of definite descriptions that rely on this knowledge as instances of *general knowledge in the larger situation*. A first-mention of *the bridesmaids* as in (6) is possible on the basis of the general speaker-hearer knowledge that weddings typically have bridesmaids (Hawkins, 1978:118).

(6) Have you seen *the bridesmaids*?

Associative Anaphoric Use. These are definite descriptions that exploit the shared knowledge of the relations between certain objects (the triggers) and their components or attributes (the associates). Whereas in larger situation uses, the trigger is the situation itself, in the associative anaphoric use the trigger is a NP introduced in the

discourse. Thus, *the author, the pages, the content* are the associates of the original trigger—i.e. a first-mention of *a book* or a subsequent mention of *the book*.

(7) I am reading *a book about Italian history*. *The author* claims that Ludovico il Moro wasn't a bad ruler. *The content* is generally interesting. (Poesio & Vieira, 1998:187)

Unfamiliar Uses. These are definite descriptions that are not anaphoric and are not associates of some trigger in the previous discourse. In other words, these definite descriptions do not rely on information about the situation of utterance. Hawkins classifies them according to their syntactic and lexical properties, as follows:

- *NP Complements*. These are definite descriptions which are characterized by the presence of a complement to the head noun.

  (a) Bill is amazed by *the fact* **that Pluto is no longer a planet**.
  (b) All the evidence pointed to *the conclusion* **that he was guilty**.

- *Nominal Modifiers*. Another form of unfamiliar definite descriptions is the presence of a nominal modifier that refers to the class to which the head noun belongs.

  (a) *The number* **nine** is my lucky number.
  (b) I don't like *the color* **red**. (Hawkins, 1978:146)

- *Referent-Establishing Relative Clauses*. A relative clause may establish a referent for the hearer without a previous mention.

  (a) *The man* **Kim went out with last night** was nasty to her.
  (b) *The supermarket* **(which is) across the street**

- *Associative Clauses*. These are the cases when modifiers of the head noun specify the set of objects with which the referent of the definite description is associated.

  (a) I remember *the beginning of the war* very well.

(b) There was a funny story on *the front page of the Guardian* this morning. (Hawkins, 1978:139)

- *Unexplanatory Modifiers*. Hawkins lists a small number of modifiers that do not fall into the above categories, but that require the use of *the*.

(a) I will never make *the same mistake* again.
(b) That was *the first time* I flew in a plane.
(c) She is *the only person* for this job.

Another line of research is based on the observation that many of the uses of definite descriptions listed by Hawkins have one property in common: the knowledge that the speaker assumes to be in the consciousness of the hearer at the time of the utterance. Simply put, the speaker (or writer) is making some assumptions about what the hearer (or reader) already knows and is able to identify the referent of the definite description. Attempts at making this intuition more precise include Christophersen's (1939) '*familiarity*' theory, Chafe's (1976) '*givenness*', Bolinger's (1977) '*knownness*', Hawkins's (1978) '*location theory*', and Prince's (1981, 1992) '*hearer-new/hearer-old information'*.

Prince expands the category of *discourse-old information*, or *givenness*, which she claims includes only entities mentioned in the previous discourse. She criticizes as too simplistic the binary distinction between given and new discourse entities that is at the basis of most previous work on familiarity, and proposes a more detailed taxonomy of 'givenness', which she refers to as *'assumed familiarity*'. According to Prince, one factor affecting the choice of a noun phrase is whether a discourse entity is old or new with respect to the hearer's knowledge. A speaker will use a definite description when he/she assumes that the hearer already knows the entity that the speaker is referring to. For example,

Nine hundred people attended *the Institute*. (Poesio & Vieira, 1998:190)

Prince argues that although the category of definite NPs usually correlates well with conveyors of hearer-old information, the correlation is not always perfect. Thus, some definite NPs can also introduce entities not assumed known to the hearer, as in:

There was *the usual crowd* at the beach.

There were *the same people* at both conferences.

For Prince, discourse-new is distinct from hearer-new. Discourse entities can be new or old with respect to the discourse model. However, for an entity, being discourse-old entails being hearer-old, but not vice-versa. For example,

(a) I need to call *a man* in California. (discourse-new/hearer-new)

(b) I need to call *a man* in California. *The man* sent me this card. (discourse-old/hearer-old)

(c) I need to call *the Institute*. (discourse-new/hearer-old)

(d) Pass me *the salt*. (discourse-new/hearer-old)

In Prince's theory, the notion of familiarity is divided into: familiarity with respect to the discourse, and familiarity with respect to the hearer. Either type of familiarity can license the use of definite descriptions: Hawkins's anaphoric uses are cases of noun phrases referring to discourse-old discourse entities as in (b), whereas, Hawkins's larger situation and immediate situation uses are cases of noun phrases referring to discourse-new, hearer-old entities as in (c) and (d).

The uses of definite descriptions that Hawkins called associative anaphoric, such as *a book … the author, the content, etc.*, are not discourse-old or even hearer-old, but they are not entirely new, either. Hawkins pointed out that the hearer is assumed to be capable of inferring the existence of these associative anaphoric NPs. Prince called these discourse entities *inferrables*, and she expanded this notion to cover another category for noun phrases that are like inferrables, but whose connection with previous hearer's knowledge is specified as part of the noun phrase itself. She referred to this category as '*containing inferrables*'. These discourse entities, as in the following example, tend to relate Hawkins's unfamiliar uses—NP complements, referent-establishing relative clauses, and associative clauses—to this category.

The door of the Bastille was painted purple. (Poesio & Vieira, 1998:190)

Taken together, as definite descriptions rely heavily on pragmatic considerations, the concept tends to be so complex that no classification scheme covers all definite

interpretations. The classification process may rely on more than just lexical cues; the writer/speaker may make assumptions about the reader/hearer's ability to use syntactic or other cues to classify a definite description. As we may see, some of Hawkins' classes of definite descriptions are specified on semantic or pragmatic grounds, others are defined in syntactic terms. Some uses tend to overlap with the other uses. For example, the unfamiliar uses with associative clauses seem related to the associative anaphoric ones, and both seem related to the uses based on referent-establishing relative clauses. From the above descriptive list, one feature that may be identified as the defining property of definite descriptions is '*uniqueness*'. Such a term is motivated by several uses of definite descriptions, particularly larger situation uses such as *the pope, the sun, the universe* and by some cases of unexplanatory modifier use, such as *the only person, the first man, the same thing*.

It should also be noted that among these classification schemes, none of them consider the so-called generic uses of definite descriptions, such as the use of *the tiger* in the generic sentence: *The tiger is a fierce animal*. This could be due to the fact that most definite descriptions are said to be 'referential'; even the first-mention NP *the salt* in: *Pass me the salt* seems to have a referential use. Thus, theories such as 'familiarity,' 'location,' or 'inclusiveness', which concern references to be identifiable, may not take generic uses into account because these uses are not 'referential'. As a matter of fact, definite uses are not always referential. For example, '*The winner will be interviewed*' can have two interpretations: the speaker may either refer to a specific person, whose identity may be mutually known to both the speaker and hearer (i.e. referential); or he/she may be referring to an activity that will take place after the winner of a competition is announced, in which case *winner* is not yet referring to any particular person.

As far as '*uniqueness*' is concerned, the generic reference of *the tiger* in the above sentence can be considered 'unique' in terms of a 'type, class, or species'. The difference is only in the physical domain (i.e. real or referential) and abstract domain of instantiation (i.e. referring to a particular class or type as a whole). In fact, when a speaker refers to one particular class or type, it can be construed as uniquely identified and thus, it is definite. Under the theory of '*uniqueness*', the notion of definiteness could

indeed cover the uses of definite-*the* and generic-*the*. As generic reference is another complex aspect, generic article uses will be reviewed in Section 2.1.2.

2.1.1.2 Indefinite Descriptions

From the explanations and discussions concerning native speaker intuition about the possible uses of the definite article, it might be assumed that the purpose of an indefinite article, in contrast, is to indicate to the hearer that the object(s) referred to does not exist in any of the speaker-hearer shared sets and that the hearer would simply have to accept that the object referred to by *a dog* or *a horse* is merely one member of an infinite class of such objects. Quirk et al. (1985) regard the indefinite article as the unmarked partner, being used (with countable nouns) when the conditions for using the definite article do not pertain.

According to Hawkins (1978), the initial theory that the objects referred to by indefinite descriptions do not exist in the speaker-hearer shared sets is not always correct. For example, if someone says *a member of parliament has resigned*, the hearer could locate the referent in the larger situation set the same way he locates *the Prime Minister*. Similarly, *pass me a bucket* could refer to one of the buckets in the immediate situation of utterance. Likewise, after mentioning *a book*, a speaker could go on to say that *a page* fell out (of that book). Thus, the referents of indefinite descriptions can be located in larger situation sets, immediate situation sets, and association sets in a similar fashion as the referents of definite descriptions. Hawkins claims that indefinites in fact can, on some occasions, refer back to objects previously introduced—i.e. to members of the previous discourse set as in:

> (1) *Some students* were standing outside the factory gate. Bill kept his eye on them. After a little while *a student* came up to him and asked him his name. (1978:174)

Notice, however, that *a student* in (1) may refer back to one of the students earlier mentioned, but equally it may not. Indefinites can thus be vague with respect to whether they are or are not genuine associates of some trigger. In general, it is the context which seems to determine how an indefinite reference is understood. Indefinite referents may be locatable in these shared sets, but whether they are or not depends on the pragmatics

of the remainder of the sentence. For Hawkins, it seems that the indefinite articles are quite neutral to the appropriate conditions of *the*. In other words, with regard to uniqueness, sentences with *a* are logically neutral to the uniqueness entailment of corresponding sentences with *the*. Hawkins (1978:203) gives examples of the ambiguities and non-ambiguities of facts such as the following:

(2) Bill didn't eat *a large cake*. [ambiguous]
(3) Bill didn't eat *the large cake*. [not ambiguous]
(4) Bill didn't eat *that large cake*. [not ambiguous]
(5) Minna wants to meet *a Norwegian*. [ambiguous]
(6) Minna wants to meet *the Norwegian*. [not ambiguous]
(7) Minna wants to meet *that Norwegian*. [not ambiguous]

From the above examples, the *specific/non-specific* ambiguity is limited only to indefinites. The explanation for this is pragmatic and involves the hearer orientation. With definites and demonstratives the hearer is pragmatically presupposed to have a form of control or knowledge of the objects referred to in a way that he is not with indefinites. For example, by using *the* in (3) or *that* in (4), the speaker is pragmatically indicating to his hearer precisely that the latter knows that there is a cake, and which cake it is that is being referred to. But by using *a* as in (2), the speaker is indicating to the hearer that the existence and identity of the object referred to does not fall under their shared experience in such a way that a definite article could be used. S*pecific/non-specific* specification is one important issue in the interpretation of indefinite descriptions; therefore, the next section will briefly review *specificity* and *non-specificity* as properties of indefinite noun phrases.

(a) Specificity

The specific sense of indefinite descriptions is found in such examples as:

(8) *A dog* bit me last night.
(9) John is marrying *an heiress*. (Werth, 1980:254)

The essential semantic feature of the specific sense is that it presupposes *existence* of its referent. Thus, sentences (8) and (9) are paraphrasable with *there BE*:

(10) There was *a dog* (which) bit me last night.

(11) There is *an heiress* (who) John is marrying. (Werth, 1980:255)

There-insertion sentences as in (10) and (11) are called *existential* sentences. Abbott (2004) noted the terms '*weak*' and '*strong*' for NPs which do and do not fit easily in existential sentences. The weak NPs are those with determiners like *a/an, some, several, many*, and the 'number determiners' (*one, two, three*, …). The strong NPs are those traditionally called *definite*—i.e. definite descriptions, demonstratives, possessives, pronouns, and NPs determined by universal quantifiers (*all, every, each*) or by *most*. Thus, sentence (10) contrasts with (12) for weak/strong distinction.

(12) *There was the dog (which) bit me last night.

Kearns (2000:120) also gives examples of specific/non-specific distinction of indefinite descriptions in:

(13) Mary wants to buy *a BMW*. She is negotiating with the owner.

(14) Mary wants to buy *a BMW*. She will look for one at the Used Car Mart downtown.

A specific reading in (13) appears where a reported thought is directed towards a particular individual the speaker refers to. *A BMW* is specific because it refers to a specific car which exists and which Mary wants to buy. Thus, using *there BE* as in: *There is a BMW that Mary wants to buy, and she is negotiating with the owner*, sounds normal. A nonspecific reading, on the other hand, appears where a reported thought is directed towards a kind or a class of objects and no particular individual is referred to. In (14), as opposed to (13), Mary has no particular BMW in mind and in fact there may or may not be such an entity. Thus, using an existential sentence in this situation is not acceptable: ?*There is a BMW that Mary wants to buy, and she will look for one at the Used Car Mart downtown*.

(b) Non-specificity

While the specific interpretation describes *something about who or what* the speaker is referring to, the non-specific interpretation describes *something about the kind of person or thing* referred to by the speaker. Thus, when saying:

(15) Melanie is looking for *a millionaire*; (Werth, 1980:257)

the speaker is not referring to any millionaire in particular. The sentence will not be interpreted as: ?*There is a millionaire who Melanie is looking for*.

In fact, definite NPs may also be non-specific:

(16) *The first man to set foot on Mars* will be a scientist.
(17) *The winner* will receive a holiday for four in Scunthorpe.
(Werth, 1980:257)

The most important distinction between the specific and non-specific senses is the presupposition of existence, which the latter cannot claim. Such a distinction has been illustrated in Werth's examples (1980:257) as in:

(18) Pass me *a sandwich*. [specific]
(19) Make me *a sandwich*. [non-specific]

Sentence (18) requests an activity which involves one of a number of existing sandwiches, whereas sentence (19) requests an activity which will bring a sandwich into existence.

From the above discussion, *uniqueness* is a primary factor in determining the definite/indefinite distinction, whereas *existence* is important when the specific/non-specific distinction is concerned in indefinite NPs. Although there are pragmatically-different interpretations with regard to specific/non-specific distinction of indefinite descriptions, it is notable that English does not mark nouns/noun phrases for specificity. Specific and non-specific singular NPs are all marked with *a* or *an*. Moreover, the non-specific interpretation tends to describe an object in a generic sense (i.e. no particular individual is being referred to). Thus, the non-specific sense of indefinite descriptions may indeed cover non-referential indefinite-*a/an* and generic-*a/an*.

So far, we have seen that the nature of the contrast between *the* and *a/an* centers on the logical property of *uniqueness* and the interpretations are carried in conversational implicatures exhibited pragmatically in *language use*. As the above-mentioned aspects are pragmatically-oriented, they can be achieved only through the introduction of the

concepts of uniqueness and speaker-hearer shared knowledge within a real model of *language use in the discourse of communication*. This is very important because as far as article instruction is concerned, most EFL lessons mainly focus on syntactic and semantic domains of English articles. Little attention is paid to pragmatic functions. This obviously explains why grammar books are inadequate to help Thai EFL learners to understand and acquire these aspects of the English language, causing article usage to pose so much difficulty even for Thai advanced learners of English.

2.1.2 Generic Reference

The specific/generic distinction is another factor which affects article choice. According to Master (1990), "the specific/generic distinction indicates when a NP is a 'real' or actual noun as opposed to when it is the idea or concept of a noun" (1990:466). Thus, the generic article refers to the article used with nouns that are of a symbolic nature (Master, 1988:18). Consequently, the distinctions of number and definiteness are neutralized because number and definiteness are not relevant for the generic concept.

Quirk et al. (1972) state that "generic reference is used to denote what is normal or typical for members of a class" (1972:147). The generic sense is commonly said to be found with all the three forms of articles, *the, a/an* and *Ø*, as in:

(1) *The horse* is a noble beast.
(2) *A horse* is a noble beast.
(3) *Horses* are noble beasts. (Werth, 1980: 252)

All three forms have readings which involve reference to a class of referents. However, many linguists agree that these different forms do not have exactly the same interpretation. Whitman (1974) argues that there is a subtle difference between generic *a/an* and generic *the*. Generic *a/an* refers to a singular example, presupposing that one entity is going to be fairly representative of all entities of the whole class. Generic *the*, on the other hand, calls forth an abstract median, the midpoint of the entire class. The generic *the* is abstract, while *a/an* is not. Whitman's proposition supports Kaluza's (1963) interpretation of generic *a/an* as referring to one typical characteristic or representative of the whole class carrying the same name, or simply "any". However generic *the*, according to Kaluza, conveys a more emphatic sense than *a/an*. In other

words, one is familiar with this class and stresses it by using 'the'. Platteau (1980) refers to the generic use with *a/an* as indefinite generics and the generic use with *the* as definite generics. For Platteau, the indefinite generic article *a/an* refers to a random element of a certain species (because the selected sample has the same default properties as all the other members of the species), while definite generic *the* gives categorical interpretation.

Cruse (2004) proposes two types of proposition involving generic reference: either something is predicated of *the whole class* referred to, or something is predicated of *each member of the class*. These two readings are known as *collective* reading and *distributed* reading. The collective reading can be clearly seen when the generic reference is indicated by a few predicates which apply to a kind as a whole, rather than to individual members of the kind. Generic-*the* strongly prefers the collective reading:

The dinosaur is extinct.
The panda may die out.
?I like watching *the panda*.

The distributed reading, in contrast, involves predicates which denote properties of individual members in the set whereby these properties can be attributed to the kind of class itself. Thus, generic-*a/an* accepts only distributive uses and cannot be combined with predicates which apply to the whole class:

**A dinosaur* is extinct.
**A panda* may die out.
A panda looks like a bear.
I like watching *a panda*.

The plural form of nouns, however, will accept either use:

Dinosaurs are extinct.
Pandas may die out.
Pandas look like bears.
I like watching *pandas*.

Master (1988) gives a detailed description of the use of the generic article. He differentiates between the *concrete* generic and the *abstract* generic. The concrete

generic refers to the representative(s) of a class and not to the class itself. Singular concrete generic nouns take the article *a/an*, and plural and uncountable concrete generic nouns take the *zero* article (*Ø*).

A dog makes a nice pet. (singular countable)
Ø Books make great gifts. (plural countable)
Ø Water is the stuff of life. (uncountable) (Master, 1988:19)

The abstract generic, on the other hand, refers to the class itself and is never concerned with representatives of that class. This applies only to singular countable nouns and is represented by the article *the*.

The alligator survives well in tropical swamps. (Master, 1988:19)

Master (1988) also identifies contexts for concrete and abstract generics. A concrete generic occurs when a generalized instance, rather than the actual noun, is meant. For example,

"What seeds a cloud without using silver iodide? *An airplane*, say researchers." (Master, 1988:19)

For this kind of context, abstract generic nouns are ungrammatical:

*Every family should own *the automobile*. (Master, 1988:20)

A concrete generic is also used in definitions:

A thermometer is an instrument that measures temperature.
(Master, 1988:19)

According to Master, definitions belong to both the abstract and concrete generic classes because in defining, the species fits into a general classification with distinguishing characteristics, whether it is an abstract generic class or a concrete generic representative or group. Definitions include attributes, classifications, comparison, etc. This accounts for the grammaticality of the following examples:

A thermometer (is an instrument that) measures temperature.

The thermometer (is an instrument that) measures temperature.
Ø Thermometers (are instruments that) measure temperature.

This context of generic use (as cited in Master, 1988:19) "has led many grammarians (e.g. Langendoen, 1970) to conclude that a generic noun can take any article with no change in meaning." In fact there are restrictions in generic use of articles in some contexts which typically call for abstract generic nouns. One important type of context which requires only abstract generic nouns is '*causality*'. Causality includes *causes* (e.g. *diagnoses, purposes*) and *effects* (e.g. *results, solutions, inventions*) as in:

The herpes virus has affected 20 million Americans. (diagnosis)
The world seems smaller because of *the telephone*. (invention)
The cyclotron ushered in the new field of particle physics. (solution/result)
(Master, 1988: 19)

Because causality is restricted only to the abstract generic, singular concrete generic nouns are not grammatical in the context of causality.

*The world seems smaller because of *a telephone*.
**A computer* is changing our society. (Master, 1988: 20)

Uncountable and plural countable generic nouns belong to the concrete generic class. NPs of this type require the *zero* article (*Ø*); for example,

Ø Milk is good for you.
Ø Pebbles are small rounded stones. (Master, 1988:20)

Notice that the forms of uncountable and plural countable genericity resemble specific uses of these nouns. The only difference between the generic and specific forms is the implied or stated use of the quantifier *some* in the surface structure of the specific use, and not with generic nouns:

Have *some milk*. (specific)
**Some milk* is good for you. (generic)
He disguised his speech by putting (*some*) *pebbles* in his mouth. (specific)
**Some pebbles* are small rounded stones. (generic)

Master (1990) lists a few constraints on generic NPs: (1) subsequent mention constraints do not apply to the generic NP; (2) generic NPs do not allow the unstressed determiner *some*, and (3) generic *a* cannot occur with nonrestrictive relative clauses. Another restriction on generic use concerns the incompatibility of generic phrases with continuous tenses. Generic NPs normally require the generality implied by the noncontinuous tenses.

The tiger lives in Asia.
**The tiger* (as a class) is living in Asia.

However, there are some continuous verbs which occur with generic phrases. These verbs share one feature—causality—which is an important characteristic of generic contexts.

The computer is becoming a fact of life.
The robot is taking the place of workers.

Master (1990) points out that generic NPs are difficult to identify because they are based entirely on context. In other words, the specific/generic interpretation can only be deduced, not from the article itself, but from the context (e.g. from the verb tense), discourse considerations and/or the nature of the sentence. For instance, *a scientist* in *John is a scientist* is specific whereas *a scientist* in *A physicist is a scientist* is generic.

From the above discussion, we can see that generic reference is still a controversial issue. Researchers have made several arguments concerning generic NPs, particularly whether the so-called generic-*a/an* should really be classified as 'generic'. According to Kaluza (1963), the interpretation of generic-*a/an* as representing 'any' one member of the whole class suggests that the representation of generic-*a/an* does not seem to be distinguishable from that of the non-specific indefinite use of *a/an*. For example,

Generic: *A horse* is a noble beast. (implying that a*ny horse* is a noble beast)
Non-specific: He needs *a pen*. (implying that any pen would serve his need)

Werth's (1980) argument supports this proposition. He maintains that only the generic phrase with *the* can be considered a true generic, for it has the overall meaning of *total* or *whole*, whereas *a/an* means 'one member', which applies to an individual in a set

rather than the whole set. Indefinite generic-*a/an* is thus not a true generic. The so-called generic-*a/an* should in fact have the same semantic representation as the *non-specific*.

In terms of generic-*the*, the representation of the whole class or species is merely symbolic. The use of *the*, in this case, implies that the speaker assumes the hearer's knowledge or sense of this symbolic representation pragmatically as generic interpretation can mostly be deduced from the context (e.g. from the verb or verb tense). Thus, generic-*the* is not so much different from definite-*the*. The difference is only in the symbolic nature of representation as opposed to a real or physical object, whereby the hearer can sense this distinction from the context. As referring to one particular type or class is always *definite* in itself in terms of *uniqueness*, it can fit in the definition of definite description uses. This is consistent with Platteau's (1980) proposition that generic-*the* is regarded as definite generics.

It should be noted that despite the complexity and subtle uses of generic descriptions, generic article usage receives minimal interest in research and pedagogy, compared to other aspects. It would be useful to examine into greater depth whether the distinction between non-referential indefinite uses and generic-*a/an* uses is important enough to classify them as two separate categories, or it is minimal enough to ignore. With regard to pedagogy, generic reference is essential for academic and technical writing (Master, 1988). Since nouns referred to in a technical sense are often depicted as a symbol rather than an actuality, a great deal of technical and academic writing is concerned with generic descriptions. Thus, it is important that further research should be conducted into this issue so that pedagogical implications can be applied to more advanced EFL article instruction.

2.1.3 Countability

Master (2002) defines a count noun as "one that has a distinct border or boundary, a 'packet' that allows the distinction of one such entity from another and thus may be counted and pluralized, e.g. pencil, star, idea" (2002:333). A noncount noun, on the other hand, is "one that has no distinct border or boundary, a mass or 'wave' that

does not allow the distinction of entities and therefore may not be counted or pluralized, e.g. plastic, flour, energy" (ibid).

Countability mainly influences article usage, especially the indefinite article *a/an* and the *zero* article (Ø). The count and mass distinction is the concept leading toward singular and plural marking. A count noun can be a single entity referred to as a singular noun and it can be morphologically marked in a plural form by an [-*s*] or [-*es*] as in *a pen* → *three pens*; *a glass* → *two glasses*. In traditional grammar books, two concepts related to the count/mass distinction are *concrete* and *abstract* nouns.

Concrete nouns are generally represented by *physical objects, substances* and *materials*, whereas abstract nouns are referred to as *qualities, feelings, states, ideas, concepts, events*, etc. (Leech & Svartvik, 1994). Concrete nouns can be seen either as count or mass. Concrete count nouns are physical objects that can be counted and can be distinguished as separate entities. Thus, they can be modified by denumerators such as numbers, and have a morphologically marked plural form: *one dog*, *two dogs*, *several dogs*. Concrete mass nouns, on the other hand, are seen as continuous entities which are regarded as substances with no natural bounds. Commonly, concrete mass nouns include materials or substances such as *water, milk, beer, gold, dirt, soap, air*, etc. Since mass nouns are not counted, they do not have singular and plural forms (but only one form) and cannot appear with other words that indicate the meaning of singular or plural—e.g. equipment, *one equipment, *two equipments, *several equipments. However, mass nouns can be modified by unspecific quantifiers such as *much* or *some*, and can be subject to division by means of certain 'gradability expressions', such as *kind, sort*, or 'partitives' such as *bar, lump, loaf, piece, pile, yard, gallon, acre*, etc.

Abstract nouns, in contrast, do not refer to real objects or entities. As abstract nouns are not visible, they can more easily be regarded as both *count* and *mass* than concrete nouns. Nouns referring to events or occasions, such as *talk, knock, shot,* or *meeting*, are usually count, while qualities and other abstract nouns tend to be mass: *beauty, honesty, wealth, happiness, progress, research, information, homework*, etc. Many abstract nouns, however, can be either mass or count: e.g. *talk, thought*. Other nouns (e.g. *difficulty, trouble, experience*) can be mass and count, but with some differences in meaning (Leech & Svartvik, 1994:44):

(1) We had little *difficulty* convincing him. [state of hardship]
(2) He is having financial *difficulties*. [problems or troubles]
(3) She is a teacher with a lot of *experience*. [knowledge or skill]
(4) Tell me about your *experiences* abroad. [past events or happenings]

Though many grammar books attempt to make a distinction between count and mass nouns, definitions for such traditional terminology, in many cases, are not of much help for learners to clearly understand the nature of the count/mass distinction. Traditional grammar views a word as being *semantically arbitrary* and therefore countability is also treated as arbitrary and is regarded as a lexical property of nouns. In fact, the notion of countability is complicated by the fact that most nouns can have their countability changed (Balwin & Bond, 2003). Alteration of countability is commonly seen in sentences such as:

(5) *beer:* There is *beer* in the glass. [mass]
(6) *beer*: There are *three beers* on the table. (packaging) [count]
(7) *carrot:* There are *three carrots* on the table. [count]
(8) *carrot*: There is *carrot* in the cake. (grinding) [mass]

(Carpenter, 1997: 329)

A term such as *beer* in (5), which is normally perceived in a mass sense, can be understood as denoting a 'package' or serving of beer, as in (6). Similarly, a term such as *carrot* in (7), which is typically countable, can be understood as a substance or mass by means of 'grinding', as in (8). In view of this complication, Wierzbicka (1988) maintains that many words can be used as either count or non-count, depending on the meaning intended. Simply put, countability judgments are based merely on "the conceptualization intended by the speaker" (p. 507). This is in line with Master's (2002) argument that although there is a greater likelihood that certain nouns will be count and others non-count, countability is largely determined by the speaker. For example, the word *pencil*, which is generally used as a count noun, can be seen in a non-count use of the word, as in "The vet found bits of chewed-up *pencil* in the dog's stomach" (p. 334). Thus, it is neither realistic nor practical to make lists of countable and uncountable nouns and to consult such lists to detect countability.

Wierzbicka (1988) proposes that grammatical countability is often motivated by the semantic distinction between object and substance reference. She gives some examples which explain her proposition. First, *size* is one crucial motivating factor which affects conceptualization for foodstuff. People always conceive of *oats* as a plural count noun and *wheat* as a mass (thus singular in form) even when these two words refer to very similar things. A few other pairs of nouns such as *rice* versus *peas* and *flour* versus *noodles* are also affected by their size. *Peas* and *noodles* are countable because individual particles of *peas* and *noodles* are much bigger than the particles of *rice* or *flour*. Thus, conceptualization of a foodstuff either as composing of individual entities or as being internally homogeneous depends crucially on how big those constituent entities are. The second important motivating factor concerns *eating habits*. *Radishes* and *olives*, which are normally countable, differ grammatically from the mass noun *garlic*, not because of the relative size, since individual olives and radishes tend to be smaller than individual heads of garlic, but because of eating habits (Wierzbicka, 1988). Olives and radishes tend to be eaten individually, one by one, whereas garlic is often chopped up and added to other foods. Grammatical difference can also be explained in terms of both the difference in size and eating habits, which accounts for differences in grammatical behavior between *onions*, which are countable, and the mass noun *garlic*. Though *onions* are not usually eaten individually, they can sometimes be cooked or pickled and served whole, whereas *garlic* is almost never served like that. Wierzbicka suggests that *vegetables* can provide clear evidence for the non-arbitrariness of grammatical differences. For example, why is *cabbage* singular whereas *Brussels sprouts* is plural? Why is *pumpkin* (as the name of a foodstuff) singular whereas *carrots* is usually plural? Her explanation is that one can see several carrots (unless it is ground up for a cake) and several Brussels sprouts on one's plate, but not several pumpkins or several cabbages. In terms of size, mass use is often motivated when the entity is either too small to attract much attention from the human consumer, such as an individual grain of rice, or too big to put on one's plate, such as an individual pumpkin or a head of cabbage.

Mufwene (1984) argues that the count/mass distinction is not lexical, and opposes the rigid labeling of nouns as either *mass* or *count* in the lexicon. He points out the difficulty in explaining what in the structure can account for the obvious linguistic-

semantic difference between the members of the pairs of words which may just appear to reflect alternative ways of conceiving of or talking about the same entities. For example,

mass	***count***	
fiction	novels	
poetry	poems	
equipment	tools	
cattle	cows	
kitchenware	kitchen utensils	
footwear	shoes	(Mufwene, 1984: 201)

Moreover, it is difficult to justify why the nouns that name objects of similar nature, such as the following, are assigned to different categories:

mass	***count***	
knowledge	beliefs	
rice	beans	
spaghetti	noodles	
garlic	onions	(Mufwene, 1984: 201)

In terms of linguistic arbitrariness, Mufwene claims that some 'mass' nouns are seen to be interpreted purely as a *linguistic* projection. For example, referents of mass nouns such as *toast, furniture, luggage, mail, livestock*, etc., are clearly perceived as divisible and do not appear to be different from count nouns such as *tables, chairs, suitcases, letters, cows*, etc. As such, Mufwene argues that '*mass-ness*' should in fact be determined by the nature of existence of a particular noun rather than linguistic arbitrariness. In this respect, he introduces the alternative terms ***individuated*** versus ***non-individuated*** to replace the traditional terms *count* versus *mass*. The term 'individuated' refers to discrete entities or sets of entities which possess a built-in mode of individuation and can be singular or plural in form. 'Non-individuated', on the other hand, is the situation when the speaker does not care to individuate the referents. In other words, the speaker uses non-individuated nouns when he wishes to refer to the objects which the nouns pick out in such a way that no division of entities into parts is presupposed. Consider the following two sentences:

(9) He likes *fruit* for dessert; or

(10) He likes *fruits* for dessert,

The choice of sentence (9) or (10) reflects the speaker's alternative perceptions of the same object(s) either as a continuous entity or a set of denumerable discontinuous objects. In most cases, how the objects are available to him, e.g. in the form of individual fruits, in slices, or as fruit salad, largely determines his particular perception of the moment and how he reproduces this in language.

In the case of liquids such as *wine, milk*, and *beer*, the built-in mode of individuation is lacking in the structure of these objects. That is, they cannot be counted as separate units. Some contexts, however, lend their own units of individuation—e.g. containers such as *glass, cup, bottle, can, carton*. Thus, it is fully understandable to say: *two coffees* to mean *two cups of coffee*, or *two beers* for *two bottles, cans*, or *glasses of beer*, depending on the availability in the situation of utterance. Some non-count nouns can be "individuated" as types, such as *wine* and *beer* in the following sentences:

(11) This is *a wine* of uncommon flavor.

(12) There is also *a beer* made of wheat rather than hops.

Allan (1980) argues that theoretically most nouns alternate between mass and count, depending on the context of use. The alterations of the count/mass distinction are the result of an individual speaker's decision to talk about or depict things one way or the other. The perceiver of his utterance will interpret the nouns in just the ways he intended for them to be; namely, as discrete sets for the count use of nouns ('individuated'—to follow Mufwene's term), or continuous ensembles for mass nouns ('non-individuated'). The conversion or alteration of countability can normally be seen in pairs of sentences such as the following (Mufwene, 1984: 202):

(a) Add a few *apples* to the tray of fruits.
Add a little *apple* to the salad. [conversion from count to mass]

(b) He drank (some) *beer*.
He drank *two beers*. [conversion from mass to count]

(c) Jane has a lot of *experience* with kids.
Jane has had a few bad *experiences* with kids. [alteration of use]

(d) Paul likes *theory*.

This is one of Paul's *theories* on mass behavior. [alteration of use]

Thus, although communication is about objects and states of affairs, it is ultimately believed to be the speaker's decision in determining (by the context of communication) how he/she wants to present them. Such a decision is made by how the speaker conceives of or perceives the objects in the context of use at the moment of speaking. The count/mass distinction, then, is not actually a distinction among words, but a distinction among ways of using those words. Countability is not a characteristic of nouns per se, but of noun phrases in the discourse of communication.

The concept of countability in English nouns is difficult for Thai EFL learners to grasp. As will be described in the following section, Thai does not mark nouns for countability and there is no need to detect whether a Thai noun is count or non-count because Thai does not have an article system. With the implied or explicit use of noun classifiers in the surface structure, all Thai nouns can be regarded as countable no matter whether their English counterparts are categorized as concrete or abstract, count or mass. The labeling of English nouns as 'countable' and 'uncountable' tends to be misleading for Thai learners. With the concept that all Thai nouns can be counted, Thai learners often have difficulties understanding why some English nouns are 'uncountable'. Their problems appear to be exacerbated when they find that some 'uncountable' nouns are sometimes used as 'countable'. As we have seen from the above discussion, it appears to be impractical to refer to English nouns as 'countable' and 'uncountable'. As most English nouns can be used as count or non-count depending on the speaker's intended meaning and the way he/she depicts things as count or mass, teaching countability should focus on the way most nouns can alternate between count and non-count use depending on the context in actual language communication.

As previously mentioned, Thai does not have an article system. Thus, the idea of the existence of articles in English may seem foreign to Thai EFL learners. This may cause them to be insensitive to the syntactic and pragmatic aspects of English articles. In fact, the concept of marking definiteness, indefiniteness, or countability is not entirely strange, for Thai uses different linguistic means to convey these notions. In the following section, inter-systemic differences between English and Thai with regard to

definiteness, genericity, and countability will be investigated. The purpose is to trace the structural, semantic, and pragmatic differences between English and Thai which could cause problems for Thai EFL learners as they may have to construct a new association between these linguistic aspects and the English article system.

2.2 Definiteness, Genericity and Countability in Thai

Although Thai does not have linguistic forms equivalent to the English articles (Iamchote, 1971; Srioutai, 2001), the concepts underlying article use—namely, definiteness, genericity, and countability—do exist in the Thai language. In what follows, I will examine how these concepts are conveyed in Thai.

2.2.1 Definiteness in Thai

Generally, traditional Thai grammars do not describe the concept of definiteness as a characteristic in the grammatical descriptions of Thai nouns. In Thai, definiteness is signaled by means of demonstratives such as *níi*, *nán*, or *nóon* (this, that, these, and those), and indefiniteness by the cardinal numeral *nὺŋ* (one) (Iamchote, 1971).

2.2.1.1 Definite Descriptions in Thai

To examine the similarities and differences in the way Thai and English marks definiteness, the following illustrates to some extent how definiteness is conveyed in Thai based on the various uses of definite descriptions in English as earlier discussed in Section 2.1.1.1.

1. Anaphoric use

Definite descriptions *the + N* which are anaphoric—i.e. pointing backward to a person or thing already mentioned—correspond to the use of *N + C (classifier) + D (demonstrative)* in Thai.

(1) I ordered a book last month, but the book has just arrived.
chăn sàŋ *nὺŋ* năŋsўy thîi lέw duən tæ̀æ x năŋsўy phˆəŋ maathўŋ
= chăn sàŋ năŋsўy lêm *nὺŋ* duən thîi lέw tæ̀æ năŋsўy lêm nán phˆəŋ maathўŋ

It should be noted that in many cases *C* or *C* + *D* in the anaphoric nominal construction in Thai can be deleted. Generally, when we already know from the context what the noun refers to, both *C* and *D,* or even *N* + *C* + *D* can be deleted. The referent is unambiguous for the hearer when the speaker says:

= chăn sàŋ năŋsўy lêm nỳŋ duən thîi lǽw tæ̀æ (năŋsўy) phˆəŋ maathўŋ

The assumption that the hearer is assumed to be able to locate the referent from the context is also applicable to other uses of English definite descriptions as illustrated in the following.

2. Associative anaphoric use

These are definite descriptions that exploit the shared knowledge of the relations between certain objects and their components or attributes (the associates), e.g. *the cover, the author, the content*, etc. are the associates of a first-mention of *a book* or a subsequent mention of *the book*. Thus, since it is clear from the context, there is no need to mark the associate noun in Thai.

(2) I bought a book. <u>The cover</u> (of the book) looks very nice.
chăn sýy nỳŋ năŋsўy x pòk duu mâak sŭəj
= chăn sýy năŋsўy maa lêm nỳŋ pòk sŭəj mâak

3. Immediate situation uses

These are definite descriptions which refer to an object in the situation of utterance. The referent may be visible, or its presence may be inferred. For example,

(3) Pass me <u>the bucket</u>.
sòŋ chăn x thăŋ
= sòŋ thăŋ hâj chăn nɔ̀j
(4) Beware of <u>the (fierce) dog</u>.
rawaŋ x x dù sùnák
= rawaŋ sùnák dù

Because the object referred to in (3), i.e. the bucket, is in the field of vision to both speaker and hearer at the time of the utterance, it is unambiguous to say the noun alone in Thai. Likewise, the inferred presence of the referent in the immediate situation in (4) allows the use of a noun without a *C* or a *C* + *D* in Thai.

4. Larger situation uses

These are definite descriptions that serve as shared knowledge among members of the same community or members in the larger situation. For example,

(5) The principal has just left for the temple.
x khruujàj phˆəŋ ja paj x wát
= khruujàj phˆəŋ ja paj wát

In (5) it is the shared knowledge that *the principal* and *the temple* refer to a particular person and place known among the community members. Thus, in Thai the *C* + *D* can be deleted.

5. Unfamiliar uses

These are definite descriptions that rely on syntactic constructions such as words, phrases, or clauses used to modify the NP. Because the modifier makes it clear what the person or thing in question refers to, only the noun alone is unambiguous in Thai.

(6) The book *that I bought* is very interesting.
x nǎŋsɤ̌y thîi chăn sýymaa x mâak sanùk
= nǎŋsɤ̌y thîi chăn sýymaa sanùk mâak

To sum up, the concept of definiteness in Thai nouns depends largely on the context of utterance. Generally, it is not necessary to mark overt forms to express definiteness in Thai.

2.2.1.2 Indefinite Descriptions in Thai

For indefinite descriptions, an indefinite NP can be used in either the specific sense or nonspecific sense. When *a* is used in a specific sense, the word *nɏ̀ŋ* is often used as the equivalent of the English NP *a + N*. The English NP with *a* corresponds to *N + C + nɏ̀ŋ* in Thai.

(7) I met a very beautiful woman.
chăn jəə nɏ̀ŋ mâak sŭəj phûujĭŋ
= chăn jəə phûujĭŋ khon nɏ̀ŋ sŭəj mâak

(8) A man came to see you.
nɏ̀ŋ phûuchaaj maa hăa khun
= phûuchaaj khon nɏ̀ŋ maa hăa khun

It should be noted that the *C + nɏ̀ŋ* in this construction is not usually deleted, although Thai speakers sometimes do. However, the use of *C + nɏ̀ŋ* helps to make the meaning unambiguous for the hearer in terms of number.

In English, the specific sense of indefinite NPs presupposes *existence* of its reference. Thus, sentence (8) can be paraphrased with *there BE* as in (9). Likewise, Thai also has another equivalent for sentence (8) which corresponds to *there BE* in English as in (9), for example,

(9) There was a man (who came) to see you.
mii nɏ̀ŋ phûuchaaj thîi maa hăa khun
= mii phûuchaaj khon nɏ̀ŋ maa hăa khun

When the indefinite article is used as a non-specifying article, the NP *a + N* also corresponds to *N + C + nɏ̀ŋ* in Thai. In this construction, the *C + nɏ̀ŋ* can be deleted.

(10) Can I borrow a pen?
săamâad chăn khŏojyym nɏ̀ŋ pàakkaa
khŏojyym pàakkaa dâam nɏ̀ŋ = khŏojyym pàakkaa nɔ̀j

(11) He is a teacher.
khăw pen nɏ̀ŋ khruu
khăw pen khruu khon nɏ̀ŋ = khăw pen khruu

(12) I want to buy a new TV.
chăn jàak sýy nỳŋ màj theewee
chăn jàak sýy theewee màj khrŷəŋ nỳŋ = chăn jàak sýy theewee màj

It can be seen that even though the indefinite article *a* generally corresponds to *nỳŋ* in Thai, there are many cases in which *a* occurs in English but *nỳŋ* does not occur in Thai. In these cases (e.g. *a* + *predicative noun*, as in 'He is a teacher'), we might say that there is no corresponding form for the article *a* in Thai.

From the way definiteness is conveyed, it can be seen that the syntactic structures of the Thai language allow for more flexibility in applying the concept of definiteness in language use, resulting in a gap of information for Thai learners of English in understanding how native English speakers associate the concept in their actual use with grammatical forms.

2.2.2 Generic Uses of Thai Noun Phrases

The generic uses of *a* and *the* in English have no corresponding form in Thai. While English requires the use of either *a* or *the*, Thai conveys genericity by the noun itself used in the discourse of communication. For example,

(13) The cow is a useful animal.
x wua pen nỳŋ miipràjòot sàt
= wua pen sàt miipràjòot

(14) A cat is a mammal.
nỳŋ mææw pen nỳŋ sàt líaŋ lûuk dûaj nom
= mææw pen sàt líaŋ lûuk dûaj nom

2.2.3 Countability in Thai

Nouns in Thai are morphologically unmarked. The grammatical descriptions of Thai nouns do not accommodate the count/mass distinction. There is no change in grammatical forms of nouns to indicate singular or plural meanings. In fact, *number* is one category in describing characteristics of Thai nouns—i.e. *singular*, *plural*, or *unidentified*. However, the singular/plural distinction of most Thai nouns depends largely on the context, or the meaning of other words appended in the sentence. The

noun classifier is considered an important device to mark number in Thai nouns. The common construction is the use of *N + cardinal numeral + classifier*. In English, liquids and substances are usually considered mass nouns and solid objects are likely to be count nouns. In contrast, by incorporating a classifier, Thai nouns, regardless of whether they are physical objects, liquids or substances, are considered to be countable. For liquid and substance, a container or a unit of measurement is often used as a counting marker, whereas for a physical object, a particular classifier is used; for example,

(15) three glasses of water
sǎam kæˆæw (C) náam
= náam sǎam kæˆæw (C)

(16) four people
sìi khon
= khon sìi khon (C)

Sometimes, plurality in Thai is marked by reduplication of noun such as:

(17) children
dèk dèk

(18) kids
nŭu nŭu

Another way to indicate plurality in Thai is by addition of other words such as:

(19) many/a number of children
mâak jamnuən dèk
= dèk jamnuən mâak

Generally, most classifier words reflect the shape, the size, or the nature of the noun they are assigned to and are usually different words from the head nouns to which they are attached, as in (15) above. Some classifiers, however, are identical to their heads, as shown in the noun phrase equivalent to *four people* in (16). Most nouns have a restricted classifier which indicates ontological structure of each entity, while some nouns require a non-restricted classifier that does not reflect the shape or size. Although it is important that classifiers are used properly in order to sound natural, in present-day

Thai, many traditional classifiers are not familiar to younger Thai people and they tend to use only a few familiar non-restricted classifiers with almost all Thai nouns.

As noun classifiers usually conform to the perception of how a particular object looks, the use of each classifier and its head tends to be semantically motivated. In other words, the characteristic of the head noun determines its designator. Thai nouns (regardless of their physical characteristics: solid, liquid, or substance) are often thought of as countable either by perception of the boundary of objects as individual entities, by containers they come in, or by their corresponding classifiers.

Another type of noun which is considered to have inherent semantic number in the lexical meaning is the *collective noun*. There are two main types of collective nouns. The first type refers to nouns used for naming a place or an organization. The focus is usually on the people in the organization rather than the organization itself. With the inherent semantic properties, nouns of this type are plural by nature. Examples are Thai words corresponding to English nouns such as *government, school, company, court of law*. The second type of collective noun consists of words representing a group of people, animals, or things. This type of collective nouns usually consists of the same set of words found in the category of noun classifier. Examples are words corresponding to English nouns such as *group* or *herd*. These nouns are different from classifiers in that they can function as a head noun in a noun phrase, whereas classifiers are appended to another noun to function as a numerical designator.

Another way of marking countability can be achieved through the use of verbs or expressions that presuppose the involvement of multiple entities; for example, words that are equivalent to the English nouns: *quarrel, fight, together, each other*.

As earlier mentioned, all nouns in Thai can be counted. In Thai, there is no need to check whether a noun is count or noncount, or whether it is singular or plural. Generally, the concept of countability in Thai nouns depends largely on the context of utterance. In other words, when it is clear from the context, countability is usually conveyed in the noun itself within the context of communication. Thus, it seems to be very difficult for Thai students to distinguish precisely between the count and non-count uses of nouns in English.

As we have seen, the structural, semantic, and pragmatic differences between English and Thai require Thai EFL learners to construct a new association between the concepts and realizations that convey definiteness, genericity, and countability in Thai NPs and the use of the English article system in NPs in English.

In the following two sections, I will review the literature on article acquisition in L1 and L2 learners. The purpose is to see whether proposals concerning L1 acquisition would contribute to the understanding of L2 acquisition and how the evidence from L1 learners could help us interpret L2 acquisition data.

2.3 First Language (L1) Acquisition of English Articles

In the literature, both L1 and L2 article acquisition research traditionally begins by investigating certain underlying features that distinguish NP environments, identifying contexts for the appearance of articles, and then examining learners' use of articles in those environments. One major theoretical approach to research on article acquisition that has been widely used to explain variation in many first and second language article acquisition studies is Huebner's (1979, 1983) noun phrase classification model, which itself was based on Bickerton's (1981) semantic wheel for noun phrase (NP) reference. Bickerton proposes two universals of NP reference: a semantic universal and a discourse universal. The semantic universal (represented by a binary feature [±SR], which stands for 'specific referent') is "that speakers indicate whether a NP refers to a specific entity or whether its reference is non-specific" (cited in Young, 1996:138). The discourse universal (represented by a binary feature [±HK], which stands for 'hearer known' or 'assumed known to the hearer') is "that speakers indicate whether they assume that the referent of a NP is known to the hearer or, alternatively, unknown to the hearer" (ibid). The semantic and discourse universal framework holds that all NPs used in discourse in any language can be classified as one of the four combinations of basic NP contexts denoted by these two binary features of referentiality: [-SR, +HK] generics, [+SR, +HK] referential definites, [+SR, -HK] referential indefinites, and [-SR, -HK] non-referentials. Huebner adopted Bickerton's semantic wheel model (Please refer to Huebner's Noun Phrase Classifications in Chapter 3) and called the four basic NP contexts as "*semantic types*". These four *types* permit us to assign a "semantic function" to each of the NP in discourse, which in turn determines article use. In fact, the

functions of the NP classified by the features SR and HK are both semantic and pragmatic, for they concern the relationships between linguistic forms and the users of those forms—i.e., the speaker and the hearer—in the discourse of communication. The meaning of a NP in discourse often extends beyond the intrinsic meaning of the NP in question. In other words, it results from the interaction between the form, the users, and the context in which it is uttered. Accordingly, the four NP types should rather be referred to as '*Pragmatic-Semantic Types*'.

In L1 article acquisition, there is strong empirical support that L1 children acquire the article system at an early age (Emslie & Stevenson, 1981; Garton, 1983; Maratsos, 1976, as cited in Cziko, 1986), and they generally exhibit a low frequency of overall errors. Cziko's (1986) study of L1 article acquisition follows Huebner's (1983) NP classification model, and proposes a four-stage sequence in L1 children's acquisition of articles based on Huebner's four NP types as summarized in Table 2.1 below.

Table 2.1: *Proposed Four Stages in the L1 Acquisition of English Articles*

	Stage 1	***Stage 2***	***Stage 3***	***Stage 4***
[-SR +HK] Generics	**Ø*	*a*	*a*	*Ø, a, the*
[-SR -HK] Nonreferential nouns attributive indefinites nonreferential indefinites	**Ø*	*a*	*a*	*a*
[+SR -HK] Referential indefinites first-mention nouns	*a,*the*	**the*	*a, *the*	*a*
[+SR +HK] Referential definites previous-mention nouns specification by entailment specification by definition unique in all contexts (etc.)	**a, the*	*the*	*(*a), the*	*the*

**Predicted errors in article use*
(Thomas, 1989:338, adapted from Cziko, 1986:881)

At Stage 1, children mark all referential nouns [+SR], both [+HK] and [-HK], with either *a* or *the*, but do not mark nonreferential nouns [-SR] with overt articles. At

Stage 2, children use *the* in [+SR] contexts and *a* in [-SR] contexts. The overgeneralization of the definite article *the* with first-mention [+SR –HK] nouns for which listeners do not have any knowledge of the reference constitutes evidence of the child's 'egocentricity'. Young children mark first-mention nouns with *the* because they assume that whatever is known to them is also known to their listener (Thomas, 1989:339). According to Cziko (1986: 896), overgeneralization of the definite article is due to the association of *the* with [+SR] environments and *a* with [-SR] environments. At Stage 3, the child begins to acquire sensitivity to the feature [±HK], possibly resulting in the reintroduction of *a* into [+SR] environments[3]. At Stage 4, the child has acquired the adult system of classification of nouns, which assigns articles according to both [±SR] and [±HK].

According to Cziko's proposed developmental sequence, children from Stage 2 onward use *a* appropriately in nonreferential [-SR –HK] contexts and *the* appropriately in referential definite [+SR +HK] environments, but they frequently use definite rather than indefinite articles with first-mention [+SR –HK] nouns. As they get older, children are more capable of adopting their audience's point of view, and consequently, the overgeneralization of the definite article decreases.

Cziko's four stages are a partial projection of Bickerton's (1981, 1984) Language Bioprogram Hypothesis (LBH), which holds that learners have an innate sensitivity to specificity/nonspecificity [±SR] of reference. Bickerton observes that creole languages produce overt articles in [+SR] contexts, and *zero*-form in [-SR] contexts. In other words, creole-speaking children use *zero* for all [-SR] nouns and use *the* for [+SR +HK] and *a* for [+SR –HK] contexts.

2.4 Second Language (L2) Acquisition of English Articles

The difficulties that L2 learners have in understanding articles are intriguing when compared with the experience of children acquiring English as their L1. There have been a considerable number of studies undertaken to examine L2 acquisition of English articles. There is a history of contrastive studies (e.g. English vs. Slavic in Kaluza, 1963) and of pedagogically-oriented analyses of the English article system (e.g.

[3] Thomas, (1989) noted, however, that Cziko's (1986) evidence that readmission of *a* at Stage 3 is unconvincing.

Grannis, 1972; McEldowney, 1977; Master, 1988, 1990, 1997, 2002). There are also descriptions of the use of English articles by learners sharing a given L1 (e.g. Kharma, 1981 for Arabic; Yamada & Matsuura, 1982 for Japanese; Agnihotri, Khanna, & Mukherjee, 1984 for Hindi and Punjabi). Among various studies, inconsistency in acquisition patterns has long been found in the literature. Brown (1983) attributes this condition to the variety of methodological approaches and disparities in data-gathering procedures. Different methodological approaches such as longitudinal (Huebner, 1983; Parrish, 1987), pseudo-longitudinal (Master, 1987), and cross-sectional studies (Yamada & Matsuura, 1982; Thomas, 1989) are considered to be a primary factor responsible for inconsistency in article acquisition orders. The longitudinal approach, though logistically very challenging to carry out, has been alleged to yield convincing results. The data obtained through cross-sectional studies are viewed to reflect only one particular phase of the interlanguage process. Furthermore, for the cross-sectional approach, which requires participants with different levels of proficiency to approximate different stages of interlanguage development, there is no shared placement standard for the participants' English proficiency.

Inconsistency may also be attributable to the sample size and varied test tasks. Generalizability of the data collected from small samples can be criticized as potentially misleading. In view of language tasks, researchers (Kharma, 1981; Mizuno, 1999; Tarone, 1985; Tarone and Parrish, 1988) find that the frequency of each error type varies depending on the nature of activities in question: whether it is a production activity such as free speaking, free writing, or it is a grammaticality judgment and other types of objective tasks. While some researchers claim that data gathered from objective tasks such as cloze tests give an inadequate view of how learners actually use articles (e.g. Thomas, 1989), others argue that production tasks may not be able to elicit article use in all environments, for learners tend to avoid using a structure if they are not sure about the usage. Avoidance of uncertain uses of articles also results in lower error rates in production tasks than in objective activities.

Another reason for inconsistency in article acquisition orders concerns the participants' different L1 backgrounds. Participants in past research represented various L1 speakers, including Japanese, Chinese, Korean, Russian, Polish, and Finnish.

Linguistic idiosyncracies underlying each L1 may have to be taken into account in direct comparison of L2 acquisition orders.

A number of L2 article acquisition studies (e.g. Tarone and Parrish, 1988; Thomas, 1989; Butler, 2002; Ekiert, 2004) have followed Huebner's (1983) NP classification, which was based on Bickerton's (1981) semantic and discourse universals, as discussed in the previous section. Huebner (1979, 1983) argues that early grammatical morpheme studies (Dulay & Burt, 1974; Bailey, Madden & Krashen, 1974; Larsen-Freeman, 1975), which had only inspected obligatory contexts, were unable to provide evidence of systematic variation in the use of a morpheme in the interlanguage development, and thus failed to give a complete picture of acquisition processes. Huebner maintains that Bickerton's semantic wheel model (as shown in Figure 3.1 in Chapter 3) allows researchers to know *when* and *how* articles are acquired. In other words, while the approach examines articles used in Standard English obligatory contexts, which tells us *when* articles are acquired with respect to one another, it also examines all occurrences of the form in question across different pre-noun contexts, which helps provide evidence of acquisition processes and systematic variation in learners' use of each article type and thus tells us *how* articles are acquired.

Huebner used his new method in a longitudinal study to investigate the use of the definite article by his participant, Ge, an adult Hmong speaker with basic-level English proficiency. Although Huebner did not conclude that Ge's learning trajectory might be universal, his findings provide evidence of systematic variability rather than random choice in article use in his subject's evolving interlanguage (Lu, 2001).

Following Huebner's longitudinal study, three comparable studies on article acquisition emerged: Parrish (1987), Master (1987), and Thomas (1989). Parrish conducted a longitudinal study of an adult beginning-level Japanese learner of English, Mari, who carried out a story-telling task every ten days for four months. Master investigated article acquisition in a pseudo-longitudinal study of 20 adult L2 learners, who were drawn from three groups of [-ART] L1 speakers (Chinese, Japanese, and Russian) and two groups of [+ART] L1 speakers (Spanish and German). To approximate a longitudinal study in a shorter period, four participants in each group of the same L1 represented four stages of interlanguage development. Master also

conducted an informal interview to elicit spontaneous speech from each participant. Thomas's cross-sectional study examined 30 adult L2 learners from 9 native language backgrounds divided into two groups: 23 participants in the [-ART] group and 7 in the [+ART] group. Each group was divided into 3 proficiency levels: low, mid, and high. The participants were paired to complete a picture-description task. Parrish, Master, and Thomas expanded their investigation to the three article types: *the, a*, and *Ø* across four NP categories. They also discussed article usage in proper nouns and idiomatic expressions, which were excluded from Bickerton's model.

Evidence from the [-ART] group in Master's study revealed that [-ART] participants used *the* considerably in [+HK] contexts, and used *a* or *Ø* to a greater extent in [-HK] contexts. The results were consistent with Huebner's finding of his participant's use of the definite article *da* for [+HK] NPs. Although Master (1997) pointed out that articles seem to be acquired differently, depending on whether or not they occur in the learner's L1, the acquisition of the definite article *the* generally precedes the acquisition of the indefinite article *a* (Huebner, 1983; Parrish, 1987; Thomas, 1989; Master, 1997), which proves to some extent that learners across proficiency levels show some sensitivity to the [±HK] distinction. However, the definite article *the* was found to be overgeneralized to both [+HK] and [-HK]. Huebner and Master referred to this dramatic rise in usage as '*the*-flooding'. *The* was found to dominate in all contexts except in [-SR, -HK] environments. Huebner and Master suggested that L2 learners might initially associate *the* with the feature of [+HK] in referential definite [+SR +HK] and generic [-SR +HK] contexts, but this did not explain why learners overuse *the* in first-mention [+SR, -HK] environment. As opposed to Huebner and Master, Parrish (1987) and Thomas (1989) claim that *the* is initially associated with [+SR] rather than [+HK] contexts. Parrish found that the use of *the* by her Japanese participant was restricted to [+SR] contexts, for *the* was markedly lacking in [-SR] contexts. Likewise, in Thomas's study, L2 participants overgeneralized *the* to [+SR] contexts, but this did not occur in [-SR] contexts. Her data showed that learners appeared to produce more instances of *the* in first-mention [+SR, -HK] contexts than in [-SR, -HK] contexts, which suggested that the learners initially associated *the* with the feature [+SR]. Moreover, Parrish and Thomas did not find *the*-flooding but overgeneralization of *Ø* in *a* and *the* contexts, while the accurate use of *a* by adult L2 learners was delayed. Thomas surmises that *Ø*-

overuse may be attributed to L1 transfer. Likewise, Master argues that overuse of *Ø* appears to result from the learners' use of *Ø* in their [-ART] L1s. At this point, Master and Thomas agreed on the effect of L1 interference of *Ø* usage. Although Master concluded that [-ART] learners tend to acquire the *zero* article first, he admitted that researchers cannot tell the difference between the *zero* article and non-use or omission of the article. In other words, acquisition is "largely by default" (p. 216). Master's data pointed out that the overuse of *zero* decreases with increasing interlanguage level, but it persists at the highest level more than overuse of the other articles.

Past research also investigated the types of errors (mostly syntactic) committed by L2 learners of different native languages (Kharma, 1981; Agnihotri, Khanna & Mukherjee, 1984), and attempted to explain the causes of some of these errors. Kharma (1981) reported that there was a substantially high percentage of errors in the use of articles by Arab university students in comparison with other types of errors, even those who had studied English for more than 12 years. However, although many kinds of errors tended to result from L1 transfer, it was not clear whether other factors had played an important role in making those errors. For instance, wrong learning strategies or tactics such as false analogy or overgeneralization or wrong equation of the two systems may also contribute to the errors. Kharma believed that inadequate teaching is the most likely factor responsible for most of the errors. Agnihotri et al (1984) pointed out in their study conducted with Indian college students that there is a vast gulf between what students need and what teachers and the available grammars offer them. They claimed that most teachers do not make distinctions clear (e.g. between the use of *a* and *an*; the use of definite and indefinite articles; or the distinction between the grammatical categories that require the use of articles and those which do not). Furthermore, students may not be instructed effectively through appropriate teaching strategies. There are certain areas which have not received sufficient remedial attention at any stage although there is evidence that shows that students' language behavior is not yet completely fossilized, such as in the area of article omission.

Taken together, early research attempted to investigate how certain underlying features that distinguish NP environments are associated with certain articles during the course of article acquisition. Considerable past research has shown that learners seem to have an innate—and therefore, universal—sensitivity to semantic and discourse

distinctions in locating reference and that there are similarities in the kinds of problems facing L2 learners, some of which may be more serious for learners from certain language backgrounds.

More recent research on L2 article acquisition also aims to provide explanations for SLA patterns in L2 English article use. Many studies have included L1 transfer and discourse universals (Jarvis, 2002; Sharma, 2005), and topic continuity (Jarvis, 2002). Considerable research reexamined earlier work; for example, Liu & Gleason (2002) reexamined Master's investigation of the overuse of *zero*, and suggested that the *zero* article and the definite article are acquired rather late. Liu & Gleason chose to examine only the various nongeneric uses of the definite article, for they claimed that the definite article has a higher frequency of use and a wider variety of usage than the indefinite article *a* or *an*.

Jarvis (2002) examined the combined perspective of two discourse universals related to topic continuity: the tendency to mark the distinction between topic (T) and comment (C) NP referents; and the tendency to mark the distinction between new (N), current (C), and known (K) NP referents. The results reveal that Finnish [-ART]-speaking and Swedish [+ART]-speaking ESL learners are sensitive to the distinction between new and not-new NP referents, but the degree to which they mark the distinction depends heavily on their L1 background. The distinction between current and known NP referents is somewhat less straightforward. Jarvis's study casts some doubt on learners' sensitivity to the topic-comment distinction. The results show little compelling evidence of either learners or native speakers differentiating overtly between topic and comment referents. Moreover, the learners' use of Ø across discourse contexts in both the Finns' and Swedes' L2 data resemble L1 patterns, which suggests that the learners' use of Ø is perhaps intentional or at least rule-governed, and may not be interpreted simply as resulting from carelessness or ignorance. Jarvis's qualitative results suggest that L2 article use is determined by several factors in addition to the discourse distinctions, and individual learners may simultaneously form multiple hypotheses about article use.

Sharma's (2005) study also evaluated the relative importance of L1 transfer and linguistic universals in the development of new principles for article use in a stable

nonnative variety of English (i.e. Indian English). The analysis reveals that the new article system is not found to be identical to the L1 article system. Though L1 transfer appears to be operative when an overt form (the specific indefinite article) exists in the L1, when no definite article exists in the L1, speakers do not completely omit the definite article in their L2 English. As Jarvis (2002) noted, the marking of broad discourse distinctions in L2 English is often triggered by the existence of differences between the L1 and L2 systems. Sharma concluded that language transfer and universals might enter into complementary partnerships, rather than acting as opposing forces. The existence of an overt specific article in individuals' L1s appears to have a strong influence on their use of the English indefinite article, but their L1 lack of a definite article permits the intervention of other universally available pragmatic principles for the use of English articles, reserving them mainly for new (or inferable) information and omitting them in more redundant contexts.

Semantic and discourse universals are still receiving considerable attention in recent research on L2 article acquisition (e.g. Lu, 2001; Butler, 2002; Ekiert, 2004). Lu (2001) adopted Bickerton's (1981) and Huebner's (1983) semantic wheel model for the analysis of NP contexts in investigating acquisition orders and underlying processes in terms of article accuracy and use by Chinese learners. The measures employed for data analysis were SOC (Supplied in Obligatory Contexts), TLU (Target-Like Use), and UOC (Used in Obligatory Contexts). Lu claimed that none of the previous studies employed TLU although it appears to be the best accuracy measure: "Parrish (1987), Thomas (1989), and Yamada and Matsuura (1982) used SOC only, and Master (1987) used SOC and UOC, but not TLU" (p. 71). Lu claimed that although SOC helps identify the acquisition orders in terms of article accuracy, it has been criticized for its failure to consider over-suppliances of a morpheme in non-obligatory contexts. If the morpheme is oversupplied or overgeneralized, SOC will overestimate the learner's accuracy. The TLU measure was designed to redress this potential inflation of SOC by taking into account the number of suppliances in non-obligatory contexts. Like TLU, UOC also takes into consideration suppliances in non-obligatory contexts. It is then able to indicate overuse or underuse of the morpheme and thus helps interpret the acquisition processes underlying the orders in terms of article use.

In Lu's study, the Mandarin Chinese [-ART] paticipants were found to have difficulty distinguishing [±HK] (e.g. misuse of *the* for *a* or Ø) and [±countability] (e.g. misuse of *a* for Ø, or Ø for *a*). Master (1987, 1997) pointed out that [-ART] speakers fail to use *a* or Ø correctly because they have difficulty judging the countability of noun phrases. Yoon (1993) also found that Japanese [-ART] learners had trouble with the article choice between *a* and Ø due to insufficient knowledge of countability. In addition, in Bickerton's semantic wheel model, *a* and Ø (as in Type 3, and Type 4) share the same contexts, but they differ in the feature [±countability]. Lu's findings also show that SOC reveals an order of *the* = *a* > Ø across proficiency levels. That is, the obligatory use of *the* or *a* is acquired earlier than Ø. The TLU measure reveals an acquisition order of *the* > *a* > Ø across proficiency groups, which means that the use of *the* is more targetlike than *a*, which is more targetlike than Ø. UOC indicates that TLU is a more reliable acquisition measure and SOC serves as a better index of accuracy level. In addition, UOC also reveals that Ø goes through a flooding-then-trickling process, *the* experiences a U-shape behavior highlighted by an overgeneralization stage, and *a* follows *the* by undergoing U-shape development as well.

Butler (2002) analyzed metalinguistic knowledge of the English article system that Japanese learners across proficiency levels employed when selecting articles in different NP environments. In addition to the binary features of referentiality, specific reference [±SR] and hearer's knowledge [±HK], which constitute the four basic NP contexts proposed by Bickerton (1981) and Huebner (1983), idiomatic expressions and conventional uses were classified as Type 5. The qualitative analysis reveals that lower proficiency learners were strongly influenced by rules they believed were given by teachers or textbooks, whereas higher proficiency learners relied more on dynamic, context-based conceptions with regard to HK and SR. Performance on the test showed that learners at higher proficiency levels had more targetlike article use than learners at lower proficiency levels, though there remained a large gap in the use of articles between the native English speakers and the most advanced Japanese learners. It was noted that nontargetlike article choices due to problems with referentiality (either misdetection of referentiality, SR and HK, or failure to consider referentiality altogether) constituted the largest percentage of errors. Based on the results, HK appeared to be more problematic than SR across groups. The higher the learners' proficiency levels were, the fewer the

number of observed problems with SR. However, problems with HK increased from the lowest ability group to the second lowest ability group and remained problematic thereafter, even among the most advanced learners. Moreover, it was found that misdetection of noun countability also constituted a major obstacle to correctly choosing articles, even for advanced learners. Those with lower proficiency tended to think that noun countability was a fixed or static entity, which they memorized from a list of count and noncount nouns from textbooks, and thus showed a lack of understanding of the dynamic ways in which SR, HK, and countability influence the proper selection of articles. Detecting HK requires accurate detection of noun countability. This determination was found to be very problematic for the learners across proficiency groups. The interview data revealed that the learners themselves acknowledged that detecting countability and referentiality of NPs presented a number of hurdles for them in accurately using English articles. Butler suggested that countability could be one of the most significant problems for learners to overcome in properly detecting HK and using articles appropriately. Thus, instruction in countability in a systematic way should be useful for learners to experience how native speakers change their perception of an entity depending on the context.

Ekiert (2004) examined the second language developmental sequence of article acquisition by adult Polish [-ART] learners in two different settings: English as a Second Language (ESL), and English as a Foreign Language (EFL). Like Butler's (2002) study, a slightly modified version of Huebner's (1983) model was employed with the addition of Type 5 (idiomatic expressions and conventional uses). The results reveal that the five NP-type uses of *a, the*, and *zero* present different levels of difficulty for L2 learners and do not appear to be acquired at the same time. ESL and EFL learners seem to follow the same path of acquisition. There is evidence supporting participants' early and accurate control of *a* in nonreferential contexts (Type 4, [-SR, -HK]). Low-ability learners also acquired *a* in first mention environments (Type 3, [+SR, -HK]) as the second article. The data show a greater rate of accuracy for generic (Type 1, [-SR, +HK]) and idiomatic (Type 5) use in low-ability learners' performance. However, as half of the items in Type 1 and Type 5 called for the *zero* article, the accuracy gained could largely be "by default" since, from the analysis, the low-ability learners had the highest rates of *zero* overuse (or failure to use any article). The data also demonstrate that while the learners used

nonreferentials (Type 4, [-SR, -HK]), referential indefinites (Type 3, [+SR, -HK]), and referential definites (Type 2, [+SR, +HK]) more accurately as their proficiency increased, their grasp of correct usage of Type 1 and Type 5 articles appeared to have regressed, which suggests that Type 1 and Type 5 would require a skillful use of *a*, *the*, or *zero*, and thus seem to be the last to be acquired. The fact that high-ability learners demonstrated no progress on generic or idiomatic use of articles could indicate that the phenomena of fossilization as well as a U-shape developmental curve may be responsible for the results. Ekiert's findings suggest that the Polish [-ART] learners replicated the L1 article acquisition order, in which *a* dominates at early stages, but the findings contradict the majority of L2 article acquisition studies in which *the* emerges early, and *a* rather late. However, *zero* overuse, which has been found in most studies with [-ART] learners, is also observed in Ekiert's study. Since it has been claimed that *zero* overproduction could be attributed to L1 transfer (Master, 1987; Thomas, 1989), Ekiert's findings give additional support to Jarvis's (2002) assumption that interlanguage is a natural language, and the use of *zero* by [-ART] learners is "rule-governed" and intentional, and should not be interpreted as either carelessness or ignorance.

While considerable research has been devoted to the study of the systematicity and variability of learners' article use in L2 article acquisition, researchers still attempt to seek explanations for the variability in acquisition orders and the interpretation that underlies article choice in L2-English. Previous investigations into the acquisition of English articles reveal that variation in article use is correlated with two main factors: L2 learners' English proficiency, on the one hand; and the existence or non-existence of the article system in learners' native languages, on the other. Much research has been undertaken to investigate performance of learners whose L1s lack articles, but the findings appear to be inconsistent across studies. The results from my pilot study conducted with two groups of Thai participants in 2006 contradict the findings of the majority of L2 acquisition studies (Huebner, 1983; Master, 1997; Parrish, 1987; Thomas, 1989) in which *the* emerges early, and *a* rather late. Evidence from the pilot study supports the order of acquisition found in Ekiert's (2004) study, in which *a* dominates at early stages for Polish [-ART] learners. High scores on Type 3 referential indefinites [+SR, -HK] and Type 4 nonreferentials [-SR, -HK] demonstrated a relative ease of detection of the [-HK] semantic feature by the Thai learners. As the participants were

asked to insert an article wherever they deemed it necessary without any blanks provided for the missing articles, the *zero* article (or rather the failure to use any article) was commonly found among the Thai participants in the study. Since it is difficult to tell the difference between the *zero* article use and non-use of (or failure to use) any article, it raises a question as to whether the *zero*-flooding which occurred in the data of this study should be termed '*zero* article overuse', or rather 'failure to use any article' (Thomas, 1989; Ekiert, 2004).

An obvious limitation of performance-based research is that one cannot be certain that L2 participants choose or do not choose English articles with a linguistically appropriate understanding of the use of articles. One often assumes that when learners make incorrect article choices, they fail to understand appropriate article usage due to syntactic differences between the learners' L1 and L2. This might in fact present one possibility among a number of other possible explanations for such errors. Furthermore, when learners use articles correctly, there is no guarantee that they have a proper understanding of article usage. They may use articles accurately but for inappropriate reasons or simply by guessing. Thus, without a clear understanding of learners' sources of errors, effectively teaching the article system remains a daunting proposition.

To accomplish this goal, the present study introduces metalinguistic tasks as a means to investigate the kinds of metalinguistic knowledge native speakers use in determining articles and to compare and contrast it with the metalinguistic knowledge Thai learners generate and use when they make article choices. The conceptual similarities and differences found in native speakers and Thai learners are expected to pave the way for further development of an appropriate pedagogical model and instructional materials to help Thai learners overcome their difficulties in English article acquisition.

In the final section of this chapter, I will review the theoretical background and research studies on L1 and L2 metalinguistics.

2.5 Metalinguistics

2.5.1 Metalinguistic Theory

Interest in metalinguistics stems from an increasing consensus among educators and researchers that a number of L2 learners lack linguistic accuracy when producing language. Since the arrival of the communicative approach to the teaching and learning of a second language in the late 1970s, a greater number of learners have been exposed to communicative activities to promote fluency, where the emphasis is on meaning as opposed to form or grammar (Renou, 2001). Such a view coincides with the general consensus that we learn our native tongue without any emphasis being placed on an awareness of or knowledge about grammar. However, research has shown that when L2 learners' attention is focused mainly on meaning, without any attention being paid to grammar, linguistic accuracy suffers (ibid).

Before discussing metalinguistics, drawing on contrasting characteristics of implicit linguistic knowledge and explicit linguistic knowledge may help provide useful insights into the nature and role of metalinguistic knowledge (MLK) in L2 learning and performance.

Explicit knowledge is defined as declarative knowledge that can be brought into awareness and that is potentially available for verbal report, while implicit knowledge is defined as knowledge that cannot be brought into awareness and cannot be articulated (Hulstijn, 2005; Roehr, 2005). Accordingly, explicit learning refers to situations when the learner has online awareness, formulating and testing conscious hypotheses in the course of learning, whereas implicit learning is an unconscious process of induction, resulting in intuitive knowledge that exceeds what can be expressed by learners (Hulstijn, 2005). A key issue is whether explicit L2 knowledge contributes to the development of implicit knowledge. Krashen (1981) has argued that the 'learning' of explicit knowledge has no impact on the 'acquisition' of implicit knowledge and is available for use only through monitoring. He maintains that only 'simple' grammatical rules can be 'learned'. In contrast, Sharwood-Smith (1981) has claimed that explicit knowledge can be converted into implicit knowledge through practice. Ellis (1993 as

cited in Han and Ellis, 1998) has argued that explicit knowledge can serve as a facilitator of implicit knowledge by helping learners to attend to linguistic features in the input.

Schmidt (1994, cited in Han and Ellis, 1998) considers both implicit and explicit knowledge as "the end-products of learning" (p. 4). According to Schmidt, the end-product may be either implicit or explicit knowledge or, more likely, a combination of both. Implicit knowledge is knowledge ***of*** language, whereas explicit knowledge is knowledge ***about*** the L2. Explicit knowledge, according to Han and Ellis (1998), can be broken down into 'analyzed' knowledge and 'metalanguage'. *Analyzed* knowledge refers to "knowledge about L2 items and structures of which learners are aware although not necessarily fully conscious" (Han and Ellis, 1998:5). *Metalanguage*, on the other hand, is "the language used to analyze or describe a language and is what learners are fully conscious of….Metalanguage can consist of technical or semi-technical terminology" (Han and Ellis, 1998:6). Research has shown that L2 learners vary enormously in the amount of metalanguage they learn (Green and Hecht, 1992).

To sum up, two principal criteria that distinguish implicit and explicit knowledge are *accessibility* and *awareness*. Implicit knowledge is easily accessed in tasks that call for fluent language performance. Explicit knowledge, by contrast, can be accessed only with controlled effort and, thus, is typically used in tasks that allow for careful planning and monitoring. Whereas implicit knowledge is unanalyzed, and consequently possessed without awareness, explicit knowledge is analyzed, and thus represents consciously held insights about language (Han and Ellis, 1998).

Metalinguistic knowledge (MLK), according to Roehr (2005), is a specific type of explicit knowledge. She refers to implicit knowledge as linguistic knowledge and explicit knowledge as metalinguistic knowledge. Roehr maintains that implicit and explicit knowledge correlate positively and interact. The following summarizes Roehr's findings regarding (explicit) metalinguistic knowledge:

(1) Explicit and implicit L2 knowledge are distinguishable but interacting constructs;

(2) MLK and L2 proficiency are positively correlated (i.e. with increasing levels of L2 proficiency, learners' use of metalinguistic knowledge increases in sophistication);

(3) MLK varies in terms of the degree of specific rule formulation and the degree of complex use of metalinguistic knowledge;

(4) Use of MLK can be understood in terms of hypothesis-testing and monitoring operations;

(5) Use of MLK is associated with consistent performance and certain decisions;

(6) Use of MLK is normally associated with successful L2 performance, but does not necessarily guarantee successful L2 performance.

The concept of metalinguistics is typically discussed in a number of disciplines (e.g. linguistics, psycholinguistics, and education) to describe a spectrum of abilities and behavior of children and adults within cognitive and language development. The developmental domain links metalinguistics to language acquisition, cognitive development, and the development of literacy skills (Birdsong, 1989). Several terms have been used, in many cases interchangeably, i.e. *metalanguage, metalinguistic awareness (MLA)*, and *metalinguistic knowledge (MLK)*. In the 1950s, linguists generally employed the term "metalinguistics" to refer to metalanguage, a language composed of the entirety of words forming linguistic terminology such as "syntax, semantics, phoneme, lexeme, etc" (Gombert, 1992).

Gombert (1992:13) defines *metalanguage* as:

> the subfield of metacognition concerned with language and its use…comprising (1) activities of reflection on language and its use; (2) subjects' ability intentionally to monitor and plan their own methods of linguistic processing (in both comprehension and production). These activities and abilities may concern any aspect of language, whether phonological, syntactic, semantic or pragmatic.

Malakoff (1992:518) defines *metalinguistic awareness* (MLA) as:

> the ability to think flexibly and abstractly about language; it refers to an awareness of the formal linguistic features of language and the ability to reflect thereupon. […]. To be metalinguistically aware, then, is to know

> how to approach and solve certain problems which themselves demand certain cognitive and linguistic skills.

The definition and operationalization of the notion of *metalinguistic knowledge* (MLK) has varied across studies. In the most general terms, MLK can be defined as learners' explicit knowledge about language (e.g. Bialystok, 1979). MLK is declarative and conscious knowledge that is potentially available for verbal report (Hulstijn, 2005). Based on Hu (2002) and Ellis (2004), Roehr (2006) defines metalinguistic knowledge more specifically as:

> a learner's explicit or declarative knowledge about the syntactic, morphological, lexical, phonological, and pragmatic features of the L2. It includes explicit knowledge about categories as well as explicit knowledge about relations between categories.

Broadly speaking, metalinguistic knowledge—or, simply put,—linguistic knowledge accessible at the conscious level—appears to be derived from linguistic development and experiences in idiosyncratic ways (Chaudron, 1983). It includes statements about intuitions of grammaticality, opinions, attitudes, perceptions of the utterance, and abstract knowledge about the language, its structure, and its uses (ibid).

Sorace (1985:239) discusses two levels of metalinguistic knowledge in the analysis of interlanguage grammars. The first level is 'the internalized, abstract knowledge' of the language; that is, the mental representation of structures and relationships associated with a given interlanguage stage. The second level is 'the procedural knowledge', which is responsible for access to the internalized knowledge and therefore for the ability to perform on the basis of it. The procedural knowledge is considered to be another kind of competence that provides the connection between interlanguage knowledge and interlanguage use. According to Sorace, the internalized, abstract knowledge can be more or less analyzed and more or less accessible, depending on the learning environment and experience. One basic principle is that *all* knowledge can be retrieved and applied productively. This is the main divergence from Krashen's Monitor Theory, which assumes that 'metalinguistic knowledge (i.e. learning) cannot function as utterance-initiator in a second language' (Sorace, 1985:252). Indeed, if one

believes that formal knowledge of a foreign language has a positive function, there is reason to justify the interaction between metalinguistic knowledge and the actual use of the foreign language. Further questions will then be open for investigation as to how to exploit this potential in a lively and communication-oriented learning situation.

For the past few decades, research on language use and language acquisition has increasingly utilized methodologies from psychology and the social sciences to substantiate theoretical positions (Chaudron, 1983). Psycholinguists and educational linguists employ metalinguistic knowledge and metalinguistic judgments as indirect indicators of differences in language structure. Researchers have recognized the value of native speakers' judgments of the 'grammaticality' and 'acceptability' of utterances and used them as the primary grounds for testing linguistic theories. L2 metalinguistic abilities have also been utilized as confirmatory data for various theories of language performance and acquisition.

The interest in metalinguistic research has been mainly in testing the adequacy of descriptive grammars of a given language. The study of L2 acquisition typically raises the unique issue of how source and target language grammars interact in the interlanguage development, and what this interaction indicates about the structure and operation of linguistic awareness itself. In fact, metalinguistic knowledge and metalinguistic judgments are naturally used in the development of materials and in instruction.

L2 research (e.g. Arthur, 1980; Gass, 1983) suggests that metalinguistic knowledge (MLK) is a reflection of developing L2 competence. MLK is often measured through learners' grammaticality judgment tests (GJT), particularly those which require error correction and justification. Judgment tasks are frequently used as a means of measuring learners' internalized knowledge. Learners' ability to judge sentences as grammatically correct or not is claimed to be a relatively direct window into competence. Sorace (1985) maintains that increases in learner judgment ability are proportional to improvement in L2 proficiency, the observable manifestation of competence. However, some studies failed to show a statistically positive relationship between MLK and L2 proficiency (Alderson and Steel, 1994; Alderson et al, 1996 as cited in Renou, 2001). Seliger (1979) found no relationship between the ability of adult L2 learners of English

to explain the rule for distinguishing *a* and *an* and their ability to use this rule in performance. Han and Ellis (1998) argue that this general failure to find any clear advantage for metalinguistic knowledge should not be taken as showing that explicit knowledge is not an important component of language proficiency.

In investigating the relative effectiveness of L2 learning under explicit and implicit conditions, it is not easy to reach conclusions as to what type of knowledge is being tapped into by the instruments (e.g. tests) that have been used to measure the outcomes of learning. The kind of knowledge a test measures is important where general language proficiency is concerned. There is also good reason to believe that the time participants are given to do the test influences the type of knowledge they draw on. For example, Bialystok (1979) found that participants were better at judging grammatical sentences correctly in spontaneous situation but better at judging ungrammatical sentences if there was a time delay. She claims that "implicit knowledge is used to decide if a sentence is grammatically correct or incorrect, but further analysis of incorrect sentences requires the formal intervention of explicit linguistic knowledge" (p. 97). Although instruments other than GJTs have also been used (e.g. discrete item production tests), again, the problem that still arises is what kind of knowledge these instruments tap into. If the instruments allow participants time to plan their sentences carefully, they may tap into primarily explicit knowledge. Thus, it may not be surprising that the results of studies that have used untimed tests (e.g. DeKeyser, 1995) show a general advantage for explicit learning. According to Bialystok (1982), metalinguistic tasks which make greater demands on higher levels of analysis are those in which learners must perform the three steps of *detecting*, *correcting*, and *explaining* the detected errors. In L1 or L2, simply judging a sentence as grammatical or not reflects implicit or unconscious knowledge. It can be done without MLA; that is, without awareness of the basis of judgment. More difficult tasks which demand justification of correction or explanation of applicable rules require the learner to access and elaborate upon their linguistic knowledge, which is a reflection of metalinguistic knowledge (Bialystok, 1982; Gass, 1983, Sorace, 1985).

Although past research (e.g. Maclay and Sleator, 1960; Hill, 1961, as cited in Chaudron, 1983) has evidenced great variability in metalinguistic judgments in native speakers and cautioned that there were limitations inherent in the idiosyncratic use of

intuition, the elicitation of metalinguistic knowledge about language structure and use has provided a potentially significant application for developmental research studies. An obvious failure in any study of metalinguistic knowledge and judgments would be the acceptance of the judgments alone as evidence for testing the hypotheses. Chaudron (1983) suggests that, in addition to the participants' responses to language stimuli, metalinguistic knowledge can be used to examine individuals' general linguistic awareness—i.e. their abilities to abstract from language use. With L2 studies, MLK certainly has clear validity as a factor in language use and development. Validation is evident in the consistency in metalinguistic reasoning and test results and in high correlations between metalinguistic reasonings.

Although there have been arguments concerning the limitations of metalinguistic knowledge, many studies (Chaudron, 1983; Chaney, 1994; Wenden, 1998; Morris, 2003) have addressed the importance of the role of metalinguistic abilities in improving learners' performance of various language skills. Researchers (e.g. Yelland, et al., 1993) maintain that metalinguistic awareness can be a necessary precursor to success in language learning. Thus, there is reason to continue to investigate this use more directly, and to transfer the applications to learners as much as possible.

To sum up, MLK comprises *analyzed* explicit knowledge of the language system, on the one hand, and the *intentional application* in their language-processing activities, on the other. From the above discussion, tasks that demand a high level of analyzed MLK include explanation of applicable rules which require learners to access and elaborate upon certain aspects of their linguistic knowledge. However, as cautioned by Chaudron (1983), there may be some limitations inherent in the use of such knowledge, particularly in the idiosyncratic use of native speakers' intuition. Thus, for validity, the elicitation of MLK about language structure and use should then be incorporated with the participants' responses to language stimuli. In this study, an objective test of English articles was used to elicit responses in terms of article usage. This task required the coordination of both analyzed MLK and the intentional application process. Another task, which required a higher degree of analyzed MLK, demanded metalinguistic explanation of applicable rules for article usage.

2.5.2 Metalinguistic Research

The interest in metalinguistic knowledge has generated a considerable amount of empirical research in both L1 and L2 learning. From the definitions discussed in the previous section, I would like to refer to L1 metalinguistics as 'L1 metalinguistic awareness' as native speakers tend to be metalinguistically aware of certain linguistic aspects from their learning environment that provides abundant input. *Analyzed* knowledge is not emphasized and the application of language processes is *less intentional*. For L2 learners, linguistic knowledge is often gained through formal L2 instruction; the role of explicit knowledge in L2 learning is more emphasized, with higher levels of *analyzed* knowledge and *intentional* application of language-processing activities. Thus, I will refer to L2 metalinguistics as 'L2 metalinguistic knowledge'.

2.5.2.1 L1 Metalinguistic Awareness

Empirical research in L1 children's metalinguistic awareness has focused primarily on metalinguistic abilities which are related to the child's ability to learn to read – phonological awareness, word awareness, syntactic or grammatical awareness, and pragmatic awareness. The most appropriate place to begin an investigation of children's awareness of language is at the phonological level because it consists of the most elementary units of language. Generally, tasks that demonstrate metaphonological awareness require children to identify and segment syllables and phonemes. Research carried out on syllable segmentation has focused on studies where children have been asked to repeat a word, once normally, and then to elide a syllable. These studies have shown that, with training, children as young as 3 years old can obtain a 60% success rate in syllable manipulation tasks.

Past research in word awareness includes investigation of children's ability to define words, which is a means of establishing their conceptions of language units. Word defining has been used to measure the child's developing ability to reflect on language as an object and to differentiate between *word* and *thing*. Snow (1990) maintains that children's word awareness fluctuates and is influenced by the opportunity to practice giving definitions.

Metasyntactic awareness, which is the ability to reflect upon the internal grammatical structure of sentences, is generally assessed through tasks that require judgments of acceptability and grammaticality. Results from Gleitman and Gleitman's (1970, in Renou, 1998) study provide evidence that children as young as 2 years old have some ability to distinguish well-formed sentences from deviant ones. However, there has been some controversy over the distinction between grammaticality and acceptability judgments. Gombert (1992), for example, claims that syntactically incorrect sentences are generally accepted by youngsters in more than half the instances in the study because young children's acceptability judgments are based more on semantic factors than on syntactic factors. Results from Tunmer and Grieve's (1984, cited from Renou, 1998) study reveal that young children (2-3 years old) tend to judge sentences according to whether or not these sentences have been understood. At 4-5 years of age, children tend to base their judgments on content. Sentences that are not liked or not believed by the child are rejected. Children at the age of 6-7 begin to realize that the form of the sentence can be separated from its meaning and is therefore able to judge the sentence on purely linguistic grounds. There seems to be a general consensus that at all ages performance is better on determining grammatical than ungrammatical utterances.

Metapragmatic awareness is the ability to develop a conscious understanding of pragmatic aspects, which include rules that represent, organize and regulate the use of speech itself in all the communicative aspects of language. Wolfson (1983, in Ellis, 2004) points out that pragmatic rules generally lie below the level of conscious awareness. Hence, some parents make explicit attempts to teach children a number of standard formulas for expressive speech acts. Aspects that have been investigated include the explicitness of a verbal message (i.e. whether a message contains sufficient information for its comprehension by an addressee), sociolinguistic appropriateness, (i.e. the relationship between linguistic choice and addressee), and linguistic humor (e.g. the manipulation of ambiguity for humorous effects). Research has shown a considerable age-related difference in children's ability to monitor the parameters of the situation of communication.

Research in adult MLA in L1 has also been conducted at different linguistic levels, but the main focus has been primarily on judgments of the grammaticality or the acceptability of sentences, primarily to help formulate theories of grammar. Generally,

since linguists are members of the particular speech community, they have often used their own intuition to differentiate linguistic levels. However, attention has shifted towards the study of native speakers considered linguistically naïve (e.g. nonlinguists) in an attempt to analyze the natural language of a community.

At the phonological and word level, much research has been conducted to assess the awareness of L1 non-literate or low-level literate adults. Studies of phonological awareness include judgments of vowel similarity (e.g. bait, bet, bat; bite, bit) and the ability to add or delete a sound segment within a syllable or a word. A study of low-level literate adults showed, as is the case with young children, that dividing one-syllable words into smaller units is more difficult than dividing words of several syllables (Hamilton and Barton, 1983, in Renou, 1998). The study of adult word awareness has been closely linked with studies of literacy. Tasks include the ability to segment phrases into words.

At the syntactic level, the focus has mainly been on grammaticality/acceptability judgments in L1 adults in terms of the syntactic aspect of sentences as opposed to the semantic aspect. A number of studies have examined the ability to judge sentences across groups of L1 adults of varying levels of education or students of different fields of study.

2.5.2.2 L2 Metalinguistic Knowledge

L2 MLK has typically been operationalized as learners' ability to correct, describe, and explain L2 errors. Research in L2 child MLK has concentrated mainly on word awareness, for it can be linked to acquisition of reading skills. However, some researchers claim that L2 learning experience can also have a positive effect on phonological awareness and syntactic awareness. For adult L2 learners, much research aims to shed light on learners' knowledge of errors that stem from their L1 and to examine advanced learners' judgments as opposed to intermediate learners. Many studies compare native speakers' judgments to nonnative speakers. Pedagogically-oriented research seeks to examine learners' judgments before and after instruction and to establish the nature of the relationship between MLK and L2 proficiency. Several

studies have sought to examine the effects of time constraint on learners' accessibility to analyzed knowledge.

In terms of phonological awareness, Rubin and Turner's (1989, in Renou, 1998) studied Grade 1 French Immersion children and found that they had greater ability in segmenting phonemes in phoneme deletion tasks than their monolingual peers. Goncz and Kodzopeljic (1991, in Renou, 1998) compared syllabic and phonemic deletion abilities in children of 5-6 and found that those who had contact with a second language outperformed monolinguals.

Word awareness (metalexical and metasemantic awareness) is considered to be a link to reading acquisition. Most studies in L2 word awareness have focused on the ability to recognize the arbitrary relationship between the word and referent. For example, being aware that words are separate from the things to which they refer can enhance reading development. Piaget (1929, in Renou, 1998) studied children's ability to differentiate between words and their referents and found that the L2 learners outperformed monolinguals. Bialystok (1986, in Renou, 1998) has concluded that contact with a L2 increases learners' control of linguistic processing. Fully bilingual children perform better than partially bilingual children on tasks requiring high levels of analysis. Research suggests that children who are fluent in two languages have greater analyzed knowledge of their language than their monolingual peers.

At the syntactic level, studies show that grammatical sentences are generally recognized earlier than ungrammatical sentences. Gass (1983) carried out a study of adult ESL learners' judgments of grammaticality using written grammatical and ungrammatical sentences produced by the participants themselves or by someone else in the group. All participants (intermediate and advanced levels) recognized close to 70% of grammatical sentences, but the advanced learners were better able to judge other participants' errors than were the intermediate learners. The greatest between-group difference was found in the higher ability of the advanced group to correct errors, which Gass attributes to increasing levels of L2 proficiency being linked to greater L2 explicit knowledge.

Sorace (1985) carried out a cross-sectional study on two groups of English-speaking students of Italian using a grammaticality judgment test where students were requested to first judge sentences' grammaticality, then to correct the ungrammatical sentences, and finally to state the rule broken. The ability to articulate a rule was found to be acquired relatively late, and is theorized to be the last stage of L2 learners' MLK development.

Roehr (2006) reports on a more recent study with English-speaking learners of L2 German. A measure of MLK comprised a two-section test. The first section was a description/explanation part consisting of sentences, each of which contained one highlighted error. Learners were required to correct, describe, and explain the highlighted mistakes. This section further contained 3 short passages paraphrased in an inappropriate manner. Learners were required to describe and explain why the given paraphrases were unacceptable. This task type was used to take into account features that depend on *pragmatic* and discursive context, i.e. features which could not easily be described or explained on the basis of an isolated faulty sentence. The test was aimed to test learners' ability to implement pedagogical grammar rules. Essentially, the targeted metalinguistic description answered the question 'What form?', and the targeted metalinguistic explanation answered the question 'Why this form?' The test items included syntactic, morphological, and lexical features of the L2. The second section was aimed at measuring learners' language-analytic ability. It required learners to employ their knowledge about grammatical categories and relations between grammatical categories typically occurring in L2 pedagogical grammar. This section was modeled on the words-in-sentences test in a four-way multiple-choice task. That is, learners were required to indicate in a second sentence the appropriate part of speech which they regarded as playing an analogous grammatical role as the highlighted part. Thus, no metalinguistic labeling or use of technical terminology was needed in this section. Results revealed that the learners' L2 grammar and vocabulary correlated strongly with L2 MLK in the 4th year learners and somewhat less strongly with L2 MLK in the 1st year learners. A possible explanation was proposed that MLK may be constructed on the basis of increased L2 proficiency. Results also suggest that learners' ability to correct, describe, and explain L2 errors and their L2 language-analytic ability may constitute the same construct—i.e. L2 metalinguistic knowledge.

Overall, a number of major findings have been proposed from early research (Bialystok, 1979; Green and Hecht, 1992; Renou, 1998; Sorace, 1985). First, positive correlations exist between levels of L2 proficiency and levels of MLK. However, the correlations may vary in strength depending on different test tasks (e.g. written vs. oral and aural tests) as the application of such knowledge does often vary across tasks. Second, it is found that learners are not necessarily able to state the grammar rule they have been taught explicitly. However, this does not mean that they are less able to correct faulty L2 items instantiating the rule in question (Green and Hecht, 1992; Sorace, 1985). Many researchers (e.g. Bialystok, 1979; Green and Hecht, 1992) maintain that some rules and categories of pedagogical grammar have been acquired and are applied more successfully than others.

While discussion and studies of L2 MLK have largely focused on syntax, L2 researchers have not specifically set out to investigate MLK of L2 pragmatic features. Many instruments that have been used to investigate learners' knowledge of illocutionary acts, such as discourse completion questionnaires, are arguably more likely to tap into explicit than implicit knowledge. The question of whether implicit knowledge and explicit knowledge mutually reinforce one another is still waiting to be addressed.

CHAPTER III

THEORETICAL FRAMEWORK UNDERLYING THE EXPERIMENTAL DESIGN

In this study, English articles are classified based on a modified version of Huebner's (1983) model of noun phrase reference. Therefore, in this chapter, I would like to discuss Huebner's noun phrase classification, which will inform the research design presented in Chapter 4.

3.1 Huebner's Noun Phrase Classifications

Huebner's (1983) model is based on Bickerton's (1981) semantic wheel for noun phrase reference (as shown in Figure 3.1). The model reflects the universal system of semantic and pragmatic marking for the speaker and hearer to locate reference in connected discourse. [±SR] is whether an entity is specific or nonspecific to the speaker, whereas [±HK] is when the speaker assumes whether the hearer has presupposed knowledge of the referent of a NP in question. The framework holds that each NP in discourse should belong to one of the four combinations of basic NP contexts denoted by these two binary features of referentiality, which define what Huebner calls 'semantic types'. These four NP 'types' determine English article use.

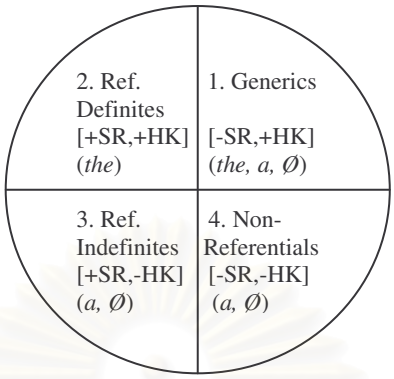
Type 1, [-SR, +HK] marks generic NPs. These nouns are marked with *the, a/an*, or *Ø*.

Type 2, [+SR, +HK] includes nouns classified as referential definites (i.e., unique, previously mentioned, physically present referents, or referents assumed common knowledge), and are marked with *the*.

Type 3, [+SR, -HK] NPs are referential indefinites, which include first-mention NPs whose referent is identifiable to the speaker but not to the listener, or NPs following existential *has/have* or *there is/are*. These nouns are marked with *a/an* or *Ø*.

Type 4 nouns, classified as [-SR, -HK] are nonreferentials, which are nonspecific for both the speaker and the hearer. These nouns are marked with *a/an* or *Ø*.

Figure 3.1 Bickerton's Semantic Wheel for Noun Phrase Reference (from Huebner, 1983)



As Figure 3.1 illustrates, Type 1 [-SR, +HK] and Type 2 [+SR, +HK] articles in the upper quadrants share a common [+HK] feature, with a distinction between [-SR] and [+SR]. Similarly, Type 3 [+SR, -HK] and Type 4 [-SR, -HK] NPs in the lower quadrants share the [-HK] feature, with a distinction between [+SR] and [-SR]. The sequencing of NP types suggests that Huebner's NPs are classified based on [±HK] as the primary feature. This is in line with Lyons's (1980) claim that articles carry a crucial pragmatic function in maintaining contextual links in the discourse of communication. [HK] is thus considered to be a more fundamental element. For ease of recalling the NP types, which will be referred to extensively in later chapters, one might conceive of the first two types (i.e. Type 1 and Type 2) as representing [+HK] NPs whose referents are assumed 'known' by the hearer, and the remaining two types (i.e. Type 3 and Type 4) as representing [-HK] NPs whose referents are assumed 'not known' by the hearer. In the semantic wheel, Type 4 [-SR, -HK] (non-referentials) adjoins Type 1 [-SR, +HK] (generics) as references of both Type 4 and Type 1 are non-specific [-SR]. This is in line with Master (1990), who argues that non-specific uses of *a/an* and *Ø* (i.e. Type 4 NPs) are eventually generic.

In the sections that follow, I would like to discuss each type of noun phrase in more detail, for the analyses of article uses presented in the next chapters are based on the articles that mark these NP types. In the following discussion, 'the speaker' represents the person who produces a statement, and 'the hearer' represents the person who receives the statement both in oral and written communication. Thus, 'the speaker' also represents 'the writer' and 'the hearer' refers to 'the reader' in the case of a written text.

3.1.1 Type 1 [-SR, +HK] Generics

Type 1 is referred to as generic NPs (Please refer to the section on *Generic Reference* in Chapter 2). The referent of a generic noun phrase appears to be nonspecific (or generic) [-SR] to the speaker, but since it refers to a class/species of entities which both the speaker and hearer share common knowledge about, the referent is assumed as known to the hearer [+HK]. These nouns are marked with *the*, *a/an,* or *Ø* and can be found in Standard English, as in the following examples:

(1) *The lion* is a beautiful animal.
(2) *A lion* is a beautiful animal.
(3) *Ø Lions* are beautiful animals.
(4) *Ø Milk* is good for you.

The structures of these generic noun phrases are:

(1) [the [+count] [+sing]] NP
(2) [a/an [+count] [+sing]] NP
(3) [Ø [+count] [-sing]] NP or [Ø [-count]] NP

Among the three articles *the, a/an* and *Ø* used in generic contexts, there is a subtle difference in terms of semantics and pragmatics, as discussed earlier in Chapter 2. In this study, generic-*the* is used to express reference of a whole class/species and generic-*a/an* is used to refer to a single member or representative of entities in a class/species.

3.1.2 Type 2 [+SR, +HK] Referential Definites

Noun phrases of this type refer to an entity whose referent is identifiable to the speaker [+SR] and is already known to the hearer [+HK]. Referential definite NPs are marked only with the definite article *the* and can be found in the following NP environments (Please see more examples in *Definite Descriptions* in Chapter 2):

(1) Unique referent: *The moon* will be full tomorrow.
(2) Referent physically present: Ask *the guy* over there.
(3) Referent assumed common/shared knowledge: Feed *the cat* before you leave.

(4) Previous-mention (Anaphoric reference):

A: So he marries a woman from England.

B: Yes, *the woman* is from London.

(5) Associative anaphoric reference:

A: I'm reading a book about Thai culture.

B: Who is *the author*?

(6) Referent specified by entailment: *The air* in this city is not very clean.

(7) Referent modified by unique adjectives (modifiers specified by definition):

That was *the first time* I flew in a plane.

I will never make *the same mistake* again.

She is *the only person* for this job.

(8) Extended reference:

A: Tom and Mary are getting divorced.

B: Yes. *The news* spread like wildfire.

The article *the* in referential definite contexts can occur with count nouns, both singular and plural in number; for example,

(1) [the [+count] [+sing]] NP: *The girl* over there is my student.

(2) [the [+count] [-sing]] NP: *The girls* over there are my students.

Since referential definite NPs mark definite or particular entities that both the speaker and hearer can identify, it is implied that such entities are perceived by the speaker and hearer as discrete sets. In other words, they are countable. As earlier discussed in Chapter 2, countability is a complicated aspect. Assigning [±count] to a noun typically depends on the speaker's intention and the surrounding context, which can be very problematic to L2 learners. Nouns that are normally perceived as noncount, such as *culture*, can be regarded as count depending on the context. Article choice is thus determined differently based on the speaker's different semantic intentions or perceptions of the object in a particular context. Countability change affects article usage, as shown in the sentences below.

(a) *Ø Western culture* places a high value on material acquisition.

(b) Working late hours for very little money seems part of *the culture* here.

As can be seen, the count/noncount distinction of a noun in most cases is not arbitrary, but semantically and/or pragmatically motivated. Therefore, one must be aware that for a NP to take *the*, the speaker has made a decision that the object in question is perceived as [+count] in a particular context of use at the moment of speaking.

3.1.3 Type 3 [+SR, -HK] Referential Indefinites

Type 3 refers to a noun phrase whose referent is identifiable to the speaker [+SR], but it is not yet a part of the hearer's knowledge [-HK]. This type includes first-mention NPs in a discourse, or NPs following existential *has/have* or *there is/are*. These nouns are marked with *a/an* or *Ø*. A count singular noun takes the indefinite article *a/an*, whereas a countable plural noun or a noncount noun takes *Ø*. Thus, the structures of Type 3 NPs are:

(1) [a/an [+count] [+sing]] NP
(2) [Ø [+count] [-sing]] NP or [Ø [-count]] NP

Examples of Type 3 NPs are:

(1) Dad gave me *a car*.
(2) He keeps sending *Ø messages* to me.
(3) Our house has *a garage*.
(4) There was *a dog* in my backyard last night.
(5) There is *Ø beer* in the glass.

3.1.4 Type 4 [-SR, -HK] Non-referentials

In Type 4, the reference of a NP is nonspecific to the speaker [-SR], nor is it identifiable by the hearer [-HK]. Type 4 NPs include nonspecific indefinites, attributive indefinites, NPs in the scope of negation, interrogation, and irrealis. These nouns are marked with *a/an* or *Ø*. When a noun has a singular form of a count noun, *a/an* is required. If a noun is plural, it takes *Ø*.

As Type 4 NPs are nonspecific to both the speaker and hearer, noncount nouns with the *zero* article are not assigned to this type. A noncount noun, by definition, is one

that has no border or boundary; thus in a nonspecific sense, it is normally perceived as generic. The structures of Type 4 NPs are:

(1) [a/an [+count] [+sing]] NP
(2) [Ø [+count] [-sing]] NP

Nonreferential NPs are found in sentences such as:

(1) I guess I should buy *a new car*.
(2) He used to be *a lawyer*.
(3) *Ø Foreigners* would come up with a better solution for this matter.
(4) I don't see *a pencil.*
(5) Do you see *a pencil*?
(6) If I had a million dollars, I'd buy *a big yacht*.

Table 3.1 below summarizes the NP structures of the four NP types.

Table 3.1: *Structures of Noun Phrase Classifications for Article Use*

Features	***Environments***	***Articles***	***NP Structures***
Type 1 [-SR, +HK]	Generic nouns	*the,* *a/an,* *Ø*	(1) [the [+count] [+sing]] NP (2) [a/an [+count] [+sing]] NP (3) [Ø [+count] [-sing]] NP (4) [Ø [-count]] NP
Type 2 [+SR +HK]	Referential definites	*the*	(1) [the [+count] [+sing]] NP (2) [the [+count] [-sing]] NP
Type 3 [+SR, -HK]	Referential indefinites	*a/an,* *Ø*	(1) [a/an [+count] [+sing]] NP (2) [Ø [+count] [-sing]] NP (3) [Ø [-count]] NP
Type 4 [-SR, -HK]	Nonreferentials	*a/an,* *Ø*	(1) [a/an [+count] [+sing]] NP (2) [Ø [+count] [-sing]] NP

As displayed in Table 3.1, it may be observed that while the pragmatic and semantic functions of the NP environments are different, there is an overlap of the syntactic structures of noun phrases across types. For example, the structure [the [+count][+sing]] can be used in Type 1 (generic) or Type 2 (referential definite) context. However, [the [+count][-sing]] can be used in only Type 2 definite context. The structure [a/an [+count] [+sing]] and [Ø [+count] [-sing]] can have Type 1 (generic),

Type 3 (referential indefinite) or Type 4 (nonreferential) interpretation. The noncount [Ø [-count]] NP can be used in Type 1 (generic) or Type 3 (referential indefinite) environment. Example sentences of article use in each NP type are shown in Table 3.2 below.

Table 3.2: *Example Sentences* of *Noun Phrase Classifications for Article Use*

Features	***Environments***	***Articles***	***Examples***
Type 1 [-SR, +HK]	Generic nouns	*the,* *a/an,* *Ø*	*The lion* is a beautiful animal. *A lion* is a beautiful animal. *Ø Lions* are beautiful animals. *Ø Milk* is good for you.
Type 2 [+SR +HK]	Unique referent Physically present Shared knowledge Previous-mention Associative use Specified by entailment Specified by definition Extended reference	*the*	*The moon* will be full tomorrow. Ask *the guy(s)* over there. Feed *the cat(s)* before you leave. I found a book. *The book* was…. I'm reading a book. *The content* is very interesting. *The air* in this city is not very clean. That was *the first time* I flew in a plane. Tom and Mary are getting divorced. *The news* spread like wildfire.
Type 3 [+SR, -HK]	First-mention nouns Existential *has/have* Existential *there is/are*	*a/an, Ø*	Dad gave me *a car*. He keeps sending *Ø messages* to me. Our house has *a garage*. There is *Ø beer* in the glass.
Type 4 [-SR, -HK]	Nonspecific indefinites Attributive indefinites NPs in negation NPs in interrogation NPs in irrealis.	*a/an, Ø*	I guess I should buy *a new car*. He used to be *a lawyer*. They are *Ø engineers.* I don't see *a pencil*. Do you see *a pencil*? If I had a million dollars, I'd buy *a yacht*.

CHAPTER IV

RESEARCH METHODOLOGY

This chapter introduces the research design and methodology as follows: the first section describes the development of research tools. In the second section, I discuss the sample groups selected for the study. The next section explains the research procedures: data collection, data analysis, and the metalinguistic knowledge coding system.

4.1 Development of Research Instruments

Tasks on English article usage were developed in response to the findings from the small-scale research project I had conducted prior to this study. Therefore, I first describe in brief the instruments used in the pilot study and some important findings that led to the development of the research instruments used in the present study.

4.1.1 Pilot Study

The pilot study was carried out to examine the pattern of article use among Thai learners of intermediate and advanced proficiency levels, with 32 Thai participants and 7 native English speakers participating in the study. The instrument used in the pilot project, adapted from Ekiert's (2004) study, consisted of 42 sentences for participants to read and insert an article wherever they deemed it necessary without any blanks provided for the missing articles. The test utilized Huebner's 4 basic types of NP classifications, with idioms and conventional uses added as Type 5. The results revealed that Type 5 was the most problematic for Thai learners of both proficiency levels. However, in the present study, Type 5 NP context was excluded for two reasons. First, article use in Type 5 NPs is normally collocational, unlikely to fall under the same syntactic, semantic and pragmatic rules as NPs in other contexts; thus they are often learned as chunks either with or without a particular article. As this study aimed to investigate rules or theories participants employed in selecting articles, Type 5 context would not be applicable if included. Second, it appears that learning to use idiomatic and formulaic expressions requires exposure to abundant input from the target language, in addition to (if indeed possible) appropriate feedback from native speakers of English to acquire usage of this article type. One might speculate that EFL learners are likely to avoid using this NP

context in their free production tasks. Since this study aimed to investigate learners' article use in an EFL setting, where there are few opportunities to practice the target language outside the classroom, Type 5 NP context was therefore discluded from analysis in the present study.

One conceivable problem with analyzing the data in the pilot study concerns what Master (1997) refers to as 'accuracy gained by default'. It was found that the most common source of errors made by the Thai participants, particularly those in the intermediate group, was the overuse of Ø. The high rate of Ø overproduction posed a problem for interpretation. Master admits that it is difficult to determine the difference between the intended use of the *zero* article and the non-use of (or failure to use) any article. The fact that the test used in the pilot study did not provide blanks for the participants to fill in an article of their choice made it difficult to identify whether or not the Thai participants' omission of an article in several places was intentional or they simply failed to use any article altogether. In the case that the article omission was not intentional, the participants' failure to use an article in contexts that required Ø could boost their accuracy rates for those particular NPs. Therefore, in the present study, blanks were provided for learners to deliberately insert *a/an*, *the*, or Ø. The purpose was to eliminate a possible 'default accuracy' effect for the *zero* article use.

Another problem with analyzing the data lies in the fact that there were a number of variable article choices found in the native speakers' responses. As the test instrument used in the pilot study was modeled on individual sentences, interpreting the pragmatic-semantic functions of articles in non-context situations could be very difficult. The decision for an appropriate article (based on the speaker's intention and the surrounding context) could vary among the native speakers. In order to understand why such variations occurred, an additional method was needed. Interviews were conducted following the article test. The interview data revealed that variations in article use were due to the native speakers' conceptual and perceptual differences of given NP contexts. Overall, their justifications for article choice were based primarily on pragmatic considerations of the NP in question. This suggested that the article test that needed to be developed for the present study should be constructed in such a way that language in discourse was provided.

In the informal interviews conducted with some Thai participants in the pilot study, it was found that they were quite capable of articulating reasons for their article choices. However, many of them had difficulty giving explanations in a face-to-face situation with the researcher. Moreover, having to give spontaneous responses without much time to reflect abstractly may hinder the participants' access of MLK. Thus, instead of incorporating interviews with the performance test, a questionnaire was developed as another tool to elicit MLK from the participants in the present study.

4.1.2 Research Instruments Used in the Study

The present study utilized two tasks on English article usage. The first task was a 'fill-in-the-article' test, which required participants' *analyzed knowledge* of article usage in completing the task. The second task demanded very high *analyzed knowledge* in supplying grammar rules or other explanations for each article choice immediately following completion of the first task. The development of each instrument is described in the following section.

4.1.2.1 Fill-in-the-Article Test

In order to test the hypothesis that article usage requires pragmatic knowledge of language in the discourse of communication, a fill-in-the-article test was developed as one task on English article usage. This test differed from a cloze test where words are deleted at regular intervals regardless of their functions. However, it was similar to a cloze test in that it requires participants to fill in deleted words, in this case English articles, in a continuous discourse. The test construction followed the procedures described below.

a) Construction of the Test

Taking Huebner's model as its base, the number of test items was initially determined so that articles could be equally distributed across the 4 NP types. As each NP type contains different articles, these 4 NP types were further sub-classified according to the number of articles that mark each type. Consequently, the test items were equally divided across the articles in each type as illustrated in Figure 4.1 below.

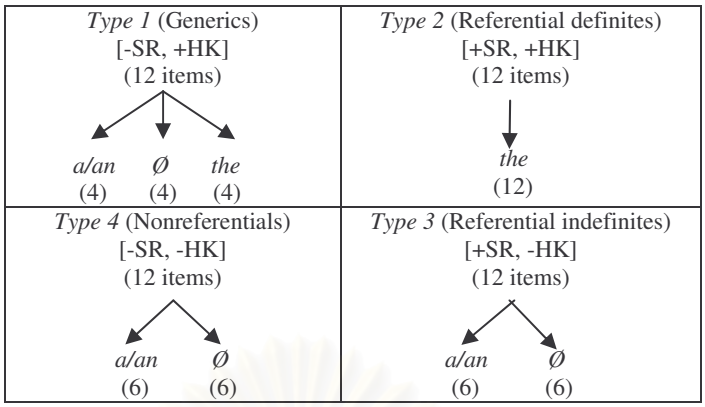


Figure 4.1: Distribution of Articles on the Fill-in-the-Article Test

Based on the design format shown in Figure 4.1, three passages from different English texts (equivalent to the level of secondary school texts) were chosen and modified. Fifty-eight articles were selected as potential test items and were deleted from these passages. The deleted articles represented the classifications in Figure 4.1, with additional items for each NP type, in case some items had to be discarded.

b) Reliability Check

The reliability of the potential test items was checked by three native English speakers, two of whom had earned a Master's degree in language education and one who had a Ph.D. in the field of linguistics and English language teaching. After an intensive training and practice session, they were asked to fill in the articles and determine their NP types, based on Huebner's framework. Disagreements over the test responses were resolved through discussion, and the passages were correspondingly modified, whereby forty-eight items were selected according to the design format. The distribution of the forty-eight items is as follows (Please refer to Appendix A):

Type 1 [-SR +HK], generics:

Ø: Items 1, 11, 13, 39
a/an: Items 16, 40, 41, 42
the: Items 3, 4, 37, 45

Type 2 [+SR +HK], referential definites:

the: Items 2, 8, 19, 22, 24, 27, 31, 35, 36, 38, 44, 48

Type 3 [+SR –HK], referential indefinites:

Ø: Items 18, 21, 25, 29, 33, 34
a/an: Items 20, 23, 26, 28, 30, 32

Type 4 [-SR –HK], non-referentials:

Ø: Items 6, 7, 10, 12, 14, 15
a/an: Items 5, 9, 17, 43, 46, 47

c) Test-Run of the Test

The test paper was prepared in duplicate. The original page was handed to the test administrator upon participants' completion of the task. The duplicate copy was retained by the participants as they were asked to perform another task to provide metalinguistic reasoning for their article choice on a questionnaire. This task will be described in the following section.

The test-run of the fill-in-the-article test was carried out with the participation of two native English speakers (1 male, 1 female), who did not have a formal language education or linguistic background. The process of the test-run was identical to the actual test administration. The participants were asked to use a blue ballpoint pen to fill in each blank by inserting one article—*a/an, the*, or *Ø*—that they thought was the most suitable. The participants were instructed that if they wanted to change any of their responses, they could do so by crossing out the article they had put in the blank and writing their new choice beside it. While performing the second task, the participants were also allowed to change their responses on the copy that they retained. However, the final article choice on their duplicate copy had to correspond to the explanation given on the metalinguistic reasoning questionnaire. The use of a ballpoint pen allowed the researcher to recognize corrections on the participant's copy. The purpose was to see if participants would want to make any change in their response when they were allowed more time to think consciously of the rules or reasons for their article choice. This was expected to reveal some important facts regarding the relationship between implicit and explicit knowledge in English article usage. The time spent by the female and male participants in the test-run of the first instrument was 5 and 8 minutes respectively.

After examining the test performance, it was found that two corrections appeared on the duplicate copy of one participant. These changes were explained to have been made in order to correct careless mistakes made when reading quickly through the text for the first time (i.e. the plural form of the given NPs had been overlooked). During the interview conducted immediately following the two tasks, the same participant decided to change another two responses due to his reconsideration of the contexts in question. He admitted that he had overgeneralized the use of *the* in those two contexts, and *Ø* should have been inserted. The test performance and the interview data from both participants served as a valuable resource for modifications of the test to eliminate potential ambiguities.

To approximate the time needed to complete the first task by Thai participants, the final version of the first instrument was field-tested with one Thai undergraduate student in the English Major Program at Dhurakij Pundit University, who was classified as an EFL learner with high proficiency level. She completed the test in 14 minutes.

Although the time to complete the test was approximated at 20 minutes, in actual test administration there was no time constraint for participants to complete the task, for setting a time limit was thought to impede low-ability learners from accessing their explicit knowledge.

4.1.2.2 Metalinguistic Knowledge (MLK) Elicitation Questionnaire

This questionnaire was developed based on an important limitation of the pilot study. As one can see, performance-based analyses merely allow researchers to assume from learners' correct responses that they acquired the article usage in question, without clear evidence as to whether their correct responses actually coincide with the linguistic rules that apply. This method might have distorted the researcher's understanding of the learners' actual knowledge and abilities in article usage. There is no guarantee that the learners chose correct articles based on appropriate reasons. In other words, the method failed to show whether learners accurately or inaccurately generalized rules to the target structures and also what rules they used (if any) when they produced non-targetlike structures. Consequently, the actual nature of learners' problems with English articles could not be identified.

The primary goal of acquisition research is to investigate where acquisition difficulty comes from; researchers accordingly have tried to explain such a difficulty psycholinguistically; for instance, as transfer from learners' L1, among other possible factors. Although one may speculate about such a possibility, there is no guarantee that the assumption is correct or that other non-speculative factors (such as transfer of training or learner's strategies) do not actually influence such a difficulty. As such, the goals of this study are to understand the metalinguistic knowledge of the English article system held by Thai students as well as to uncover the reasons why they have great difficulty with English article usage.

Another notable point was that, from the interviews conducted with the native speakers following the article test in the pilot study, there was solid evidence that native speakers had the ability to articulate reasons for their article choice. Indeed the interview data provided rich information regarding the native speakers' differing pragmatic considerations of the given contexts. Such evidence does not support a common belief that native speakers often use English articles intuitively and thus are incapable of explaining why they use certain articles in given contexts. The native speakers recruited for the present study were college students; they, therefore, should be more or less able to explain their article choices based on their learning experiences in school.

As earlier mentioned, incorporating interviews with the performance test would be problematic. Face-to-face interviews not only consume much time but also cause a great deal of anxiety for Thai students when they have to respond spontaneously in the presence of the researcher. Thus, a MLK elicitation questionnaire was developed. The participants could take as much time as they needed in giving responses without fear of embarrassment. Moreover, the technique was also less time-consuming than conducting interviews with all participants.

The questionnaire was prepared in two languages: English and Thai. (Please refer to Appendix B) The participants were asked to write in their own words and in their native language the metalinguistic explanation corresponding to each article choice in the space provided. To evoke prompt elicitation of participants' metalinguistic knowledge they employed in selecting certain articles, the questionnaire was used immediately following the completion of the article test. The reasons given by the participants may or

may not have coincided with linguistic rules that apply in a certain context. This allowed the researcher to better understand what the students knew and what they did not know about English article use than providing prompts for the participants to respond to. Prior to actual administration, both the English and Thai versions of the questionnaire were field-tested with two native English speakers and a Thai EFL learner concurrently with the fill-in-the-article test. The time spent by each participant to complete the questionnaire ranged from 35 to 45 minutes. However, like the fill-in-the-article test, this metalinguistic task was not timed in actual administration.

4.1.2.3 General Questionnaire

In addition to the two tasks described above, another type of questionnaire was also included. This questionnaire was prepared in Thai. (See Appendix C) Two main purposes for utilizing this questionnaire were: (1) to sort out Thai participants who would not be suitable for the purposes of this present study (e.g. those Thai learners who had been in a country where English is spoken natively for longer than three months); and (2) to collect background information on the students' histories of learning English. The questionnaire comprised three parts and it took approximately 10 minutes to complete. The first part required students to provide personal and contact information (e.g. their names, ages, email addresses and telephone numbers). The second part was composed of general questions regarding the students' EFL histories. The third set of questions aimed to collect information on students' experience of learning the English article system. The questionnaire was given to Thai students prior to the administration of the article test. Based on the answers to the questionnaire, no Thai students who had initially been recruited were excluded from participation in the present study.

For native English speakers, a short questionnaire was used primarily to verify that English was actually their first language and that they were not studying in the field of English, language studies and/or linguistics. Participants were also requested to provide personal and contact information. Based on the answers to the questionnaire, fourteen students were excluded from the study. This will be described in more detail in the following section.

4.2 Participants

Ninety undergraduate students participated in this study. Thirty of them were native English speakers and sixty of them were Thai students learning English as a foreign language (EFL). The recruitment of these participants is described below.

4.2.1 Native English Speakers

Thirty native English speakers were recruited from Wayland Baptist University in Plainview, Texas, U.S.A. All of them were undergraduate students in fields other than English, language studies and/or linguistics. The primary purpose was to eliminate the influence of familiarity with linguistic issues that may affect the participants' metalinguistic knowledge. These students were recruited from two randomly-selected groups of a fundamental English course, comprising a total of 44 students who were majoring in different fields. From the background questionnaire, 10 students in these two groups spoke English as a second language. Their test papers were accordingly excluded from the data analysis. Among the thirty-four native English speakers in the two groups, four students were studying in the field of English, English Language Arts, and English Education. Therefore, only thirty students were suitable for the purposes of this study. These thirty participants included ten males and twenty females, ages ranging from 18 to 23. They were designated the native speaker (NS) group.

4.2.2 Thai EFL Learners

Sixty Thai participants were first- and fourth-year undergraduate students in the English Major Program at Dhurakij Pundit University (DPU), which is a private Thai university located in Bangkok, Thailand. These nonnative speaker (NNS) students were selected and divided into two proficiency groups, *high* and *low*, based on their relative English proficiency, with thirty students placed in each group (hereafter referred to as NNS-H and NNS-L respectively). The classification was based on their scores on the DPU-TEP (DPU Test of English Proficiency) taken by 196 English Major students (128 freshmen and 68 seniors). Among these 196 students, thirty students with the highest scores and 30 students with the lowest scores were recruited to participate in the study. Out of the total DPU-TEP score of 130 (100%), the 30 high proficiency students' scores ranged from 107 (82.3%) to 118 (90.8%), whereas the scores of the 30 low proficiency

students ranged from 40 (30.8%) to 64.5 (49.6%). Although it cannot be claimed that the DPU-TEP is a standardized test, the reliability of this instrument has been reported at the Cronbach's alpha of .901. This should serve well for the purpose of this study.

None of the NNS groups both the high and low-proficiency had been in a country where English is spoken natively for longer than three consecutive months, and none had extensive exposure to English or any other languages besides Thai. A questionnaire with a list of questions regarding the learners' demographic information and their English language education background was used at the time of the participant recruitment. (See the above section on *General Questionnaire*) The high proficiency group included 7 males and 23 females, ages ranging from 20 to 26. These students had the minimum of 10 years up to a maximum of 18 years of formal English instruction. The low proficiency group comprised 4 males and 26 females, ages ranging from 17 to 22. Their length of formal English instruction ranged from 6 to 13 years.

The justification for using high and low proficiency learners is based on Oller and Redding's (1971) assumption that article usage in interlanguage relates to overall proficiency in the target language. Oller and Redding correlated accuracy in article usage on a written test with standardized ESL placement test scores and found a strong relationship between the two measures, with 78% of the variance on the article test being explained by scores on the placement test. Oller and Redding (1971:94) concluded that "the acquisition of articles usually goes hand-in-hand with other English language skills." This serves as strong grounds for developmental hypotheses that learners at varying proficiency levels perform differently in article use on the basis of their different metalinguistic knowledge. Accordingly, learners with high proficiency are expected to have language abilities and metalinguistic knowledge which are more consistent with that of native English speakers than low proficiency learners.

4.3 Data Collection

The data collection was conducted in 2 phases:

(1) Subsequent to the development of the research instruments, two groups of sixty Thai participants were recruited according to the procedure described above. These sixty students were asked to complete the EFL background questionnaire to verify their

eligibility. The article test and the MLK elicitation questionnaire were administered to both groups of Thai participants.

(2) The second phase of data collection was conducted in the United States. Non-English major undergraduate students were recruited from Wayland Baptist University, as earlier discussed. The article test and the MLK elicitation questionnaire were administered to thirty native English speaker participants.

4.4 Data Analysis

The data analysis was conducted quantitatively and qualitatively using the following procedures:

4.4.1 Quantitative Data Analysis

The quantitative analyses focused on participants' performance on the article test and metalinguistic knowledge categorized according to the coding system which will be discussed in the next section. Scores from the article test and the MLK coding were tabulated and calculated using the following statistics:

(1) percentage
(2) mean scores
(3) frequency distribution
(4) Kappa coefficient
(5) analyses of variance (ANOVA)

4.4.2 Qualitative Data Analysis

The qualitative analyses focused on the metalinguistic explanations for article use provided in the MLK elicitation questionnaire by the three sample groups.

4.5 Analysis of Metalinguistic Knowledge

4.5.1 Metalinguistic Knowledge Categories

In the metalinguistic knowledge analysis, a coding system was set up based on the actual statements given in the MLK questionnaire. Key words were listed and used

as the basis for classifying their metalinguistic reasoning into two main categories: (1) 'explicit' reasons (i.e. participants identified grammar rules or other explanations for their use of articles explicitly), and (2) 'non-explicit' reasons (i.e. participants did not explicitly identify any specific reasons for selecting certain articles). The participants' reasons that were classified as 'explicit' were further divided into 3 subcategories: *pragmatic* (P), *semantico-syntactic* (S), and the combination of *pragmatic and semantico-syntactic* (PS). The 'non-explicit' reasons were divided into 3 subcategories: *impressionistic* (I), *guessing and pseudo-guessing* (G), and those with irrelevant responses or without responses were labeled as *no reason given* (N). Figure 4.2 below illustrates all the metalinguistic knowledge categories and the coding system.

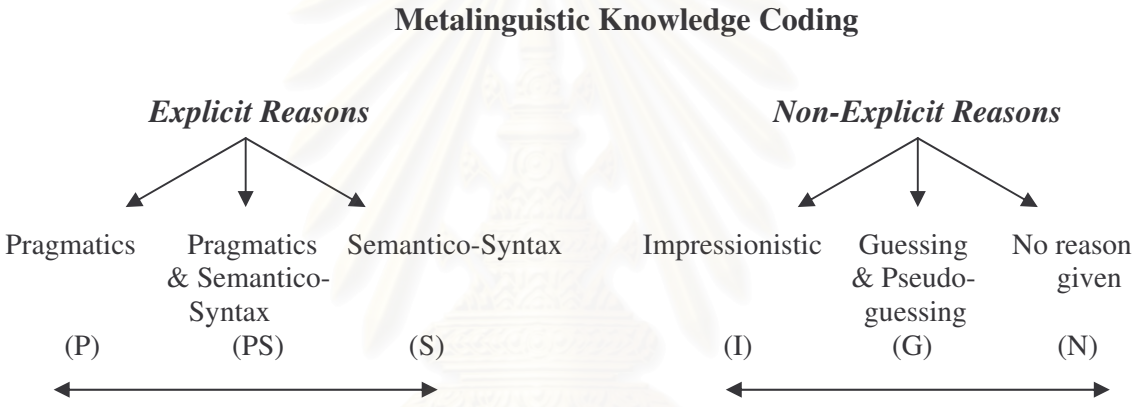


Figure 4.2: Metalinguistic Knowledge Categories and the Coding System

(1) The 'Explicit' Reasons

The 'explicit' reasons comprise 3 subcategories: pragmatic (P), semantico-syntactic (S), and the combination (PS). Key words elicited from the participants' metalinguistic reasoning for each subcategory are shown below.

Pragmatic Reasons (P)

The reasons that fall into pragmatics consist of explanations that go beyond the reference to the semantic meaning or syntactic structure of the NP in question. They are based entirely on the context or discourse of communication. Samples of key words consist of:

- (not) specify; (not) specific
- (in) general

- a particular/certain entity
- referring to the whole/type/group/class/species
- referring to a/any random entity/member of the group
- firstly introduced/not known yet
- previously mentioned/described

Semantico-Syntactic Reasons (S)

These are the explicit reasons based on the reference to the meaning of the NP in question or the grammatical structure embedded in a given NP. The following key words were described by the participants:

- singular/plural noun
- referring to one or many things (countability)
- used before a noun
- next word begins with a vowel
- rule of the following vowel
- rule of the following consonant

Pragmatic and Semantico-Syntactic Reasons (PS)

The reasons provided for this subcategory contain words classified for both pragmatic and semantico-syntactic explanations. Examples were when participants wrote the following:

- (not) specific (P) and singular/plural (S)
- one/many/multiple (S) and general/specific (P)
- (not) specific (P) and begins with a vowel (S)
- used at the beginning of the sentence (S) and talking about something in general (P)

(2) The 'Non-explicit' Reasons

The reasons classified as non-explicit were divided into 3 subcategories: impressionistic (I), guessing/pseudo-guessing (G), and no response (N). The following are key words elicited from the participants' metalinguistic reasoning for each subcategory.

Impressionistic Reasons (I)

Reasons classified as 'impressionistic' are those that cited intuition/ feelings/beliefs in using a certain article for the context in question. Examples include:

- it sounds correct/ it best fits the sentence/ it was the best choice
- any article wouldn't fit/ nothing could fit there
- I don't think you need an article/ doesn't require an article
- it's better without an article
- the sentence flows without an article
- an article not needed to make sense
- I believe this is the correct article to use
- 'the' should be the proper article to use
- sounds better than 'a' or 'an'

Guessing or Pseudo-Guessing (G)

From the explanations given by the native English speakers, it was determined that this subcategory was not found in any of their responses. Generally when native English participants could not identify any specific reason for their article choice, it was because in some instances they simply used their intuitive (I) judgment without being able to articulate the reason for that particular usage. No one indicated that they guessed at any of their article choices. Some native English participants may have changed a few responses, but their explanations seemed to suggest that they merely changed their interpretation of the NP context, or they later recognized some clues in the context that they failed to detect earlier.

For the Thai learners, the use of guessing or pseudo-guessing strategies was expressed by statements such as:

- I'm not quite sure…
- I just guessed…
- I guess this might be the best choice
- Others don't seem to work

Please note that the last example: *Others don't seem to work* was sub-categorized as 'pseudo-guessing' rather than 'impressionistic'. 'Impressionistic' reasons are classified based on the participants' "feelings" toward their choice as being the 'right' one. This strategy is different from 'elimination', which I classified as 'pseudo-guessing'. When the participants eliminate all other 'improbable' choices, it does not mean that they feel what is left for them to choose should be the correct answer. As their reason does not imply that they 'think, believe or feel' that their choice should be the 'right' one, such a decision should be a part of 'guessing'.

No Response or No Reason Given (N)

For this category, participants left the space blank or had written something that did not explain why they chose the particular article; for example, they merely repeated or translated the NP. Some learners admittedly stated that they had no idea; they just picked one article without any reason at all.

4.5.2 Inter-Rater Reliability

To attain inter-rater reliability, 4 out of 30 completed MLK questionnaires were randomly selected from each group of participants as representative samples. These randomly selected samples, representing 13.3% of participants in each group, were coded by two Thai bilingual researchers based on the above-mentioned coding system. One of the coders was the author of this study. Disagreements in coding were resolved and some clarifications were made through extensive discussion. The percentage agreement on coding attained between the two coders was 99.5% for the NS group, 99.0% for the NNS-H, and 99.5% for the NNS-L. The Kappa coefficient of the randomly selected samples was calculated and the values attained were .992, .985, and .990 for the NS, NNS-H, and NNS-L respectively. This was considered a sufficient degree of agreement between the coders for the present study. The remaining metalinguistic questionnaires were coded by the author based on the coding system agreed upon by the two raters during the inter-rater discussion.

CHAPTER V
RESULTS (1)
TEST PERFORMANCE AND METALINGUISTIC KNOWLEDGE IN ENGLISH ARTICLE USE

The results of the test performance and the metalinguistic knowledge in English article use reported in this chapter were both quantitative and qualitative in nature. The quantitative analyses focus mainly on the test scores of the three participant groups in relation to the distribution of metalinguistic knowledge given for their article choices. Metalinguistic explanations have been examined and discussed qualitatively. In order to present and discuss the findings of the study according to its objectives, in this chapter I will first describe the scores on the target article use in the native speakers and Thai learners. I then compare and contrast similarities and differences in the metalinguistic knowledge given in justifying their article choices. In the next chapter, I will present the metalinguistic knowledge the participants used for both the targetlike and non-targetlike article choices in order to identify problems in learners' article usage.

The organization of the present chapter is as follows: first, I present the overall test performance for target article use by the three sample groups, using descriptive statistics and analyses of variance (ANOVAs). Second, by using the same analyses, I report on the targetlike rates for the NP types across the groups. Third, I further examine and present the participants' performance on the eight NP sub-types (i.e. NP types marked with each target article), using the same statistical analyses. Similar analyses were conducted to investigate the participants' metalinguistic reasoning on their article choices which had been coded as described in Chapter 4. First, the analysis was based on the overall test performance across the groups, then on the performance of each NP type, and finally on the target article use in each sub-type across the groups.

5.1 Difference in Test Performance

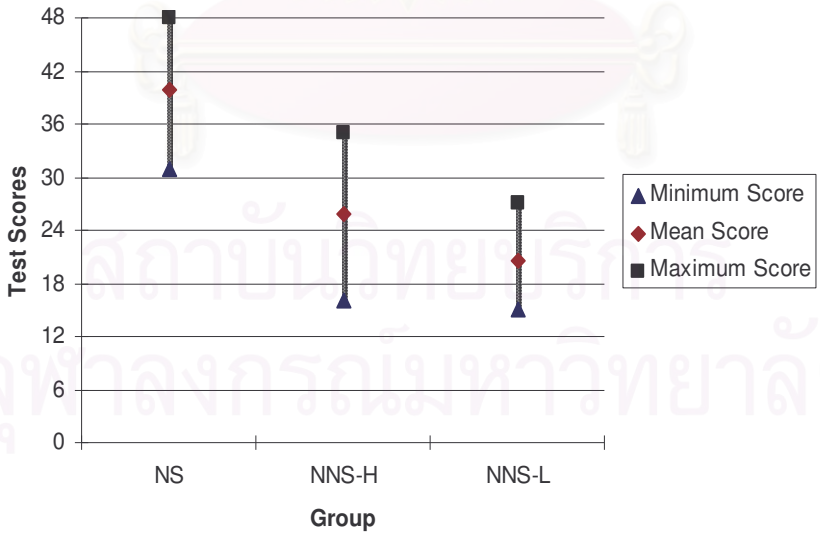
5.1.1 Difference in the Overall Test Performance

The maximum, minimum and mean of raw scores for the target article use in all NP types are displayed in Table 5.1. The same results are also illustrated in a graph in Figure 5.1.

Table 5.1: *Maximum, Minimum and Mean of Raw Scores on the Article Test by Group*

Group	Maximum Score	Minimum Score	Mean Score	Std.	Average Mean Score in Percentage
NS	48	31	39.77	4.329	82.8 %
NNS-H	35	16	25.8	4.429	53.8 %
NNS-L	27	15	20.63	3.508	43.0 %

Figure 5.1: *Maximum, Minimum and Mean of Raw Scores on the Article Test by Group*



As the fill-in-the-article test contained 48 items for all NP types, the maximum score on the test was 48 for each participant. In Table 5.1, the raw score of 48 was the highest score found in the NS participants and 15 was the lowest raw score found in the NNS-L group. The means of the test scores linearly increased in relation to proficiency

levels (i.e. 39.77 > 25.8 > 20.63). However, the gap in the mean scores between the NNS-H and NNS-L learners (i.e. 25.8 and 20.63) was relatively narrow compared with a more sizable gap between the NS and NNS-H groups (i.e. 39.77 and 25.8). The numbers in the second column in Table 5.1 show that the minimum test score of 16 in the NNS-H group was close to the lowest score of 15 obtained by the NNS-L learners. However, the maximum score of 35 in the NNS-H group was considerably higher than the highest score of 27 in the NNS-L group. A broader range of scores among the high-proficiency learners and the higher standard deviation suggest that the NNS-H group was assumed to be a more heterogeneous group than the NNS-L group. As Figure 5.1 indicates, the overlap of scores between the two groups of Thai learners was substantial. Surprisingly, there was also an overlap of scores between the NS and the NNS-H groups. The highest score of 35 obtained by the NNS-H was higher than the lowest score of 31 in the NS group. This could suggest that some high-proficiency learners outperformed some native speaker participants. Moreover, a broad range of scores (i.e. from 31 to 48) in the NS group indicated that the native speakers' performance on the test varied substantially. This is quite surprising given that the test items were constructed based on agreements among native English speakers who were experts in linguistics and ELT.

In order to see the differences in scores on the test among the three groups, a one-way ANOVA was introduced. The results showed that there were differences in the test scores among groups ($F(2, 87) = 174.057$, $p < .001$). Tukey's HSD Test further revealed that there were statistically significant differences in the mean of raw scores between native speakers and Thai learners with high and low proficiency levels, as well as between the two groups of Thai learners, at the alpha .05 level. (See Table 5.2)

Table 5.2: *Comparisons of the Mean of Raw Scores on the Article Test*

Group Comparison	Difference of Means of Raw Scores	Sig.
NS vs. NNS-H	13.97*	.000
NS vs. NNS-L	19.14*	.000
NNS-H vs. NNS-L	5.17*	.000

* The mean difference is significant at the .05 level

5.1.2 Difference in Performance on the Article Test by NP Types

Table 5.3 illustrates the frequency of score distribution and the percentage (shown in parentheses) across four NP types by group. The same percentage scores are also displayed in a line graph in Figure 5.2.

Table 5.3: *FD and Percentage of Target Article Use by Group and NP Type*

Group	NP Type			
	Type 1 (N=360)	Type 2 (N=360)	Type 3 (N=360)	Type 4 (N=360)
NS	268 (74.5%)	**347 (96.4%)**	283 (78.7%)	295 (82.0%)
NNS-H	159 (44.2%)	**263 (73.1%)**	160 (44.5%)	192 (53.1%)
NNS-L	149 (41.4%)	**194 (53.9%)**	136 (37.8%)	140 (38.9%)

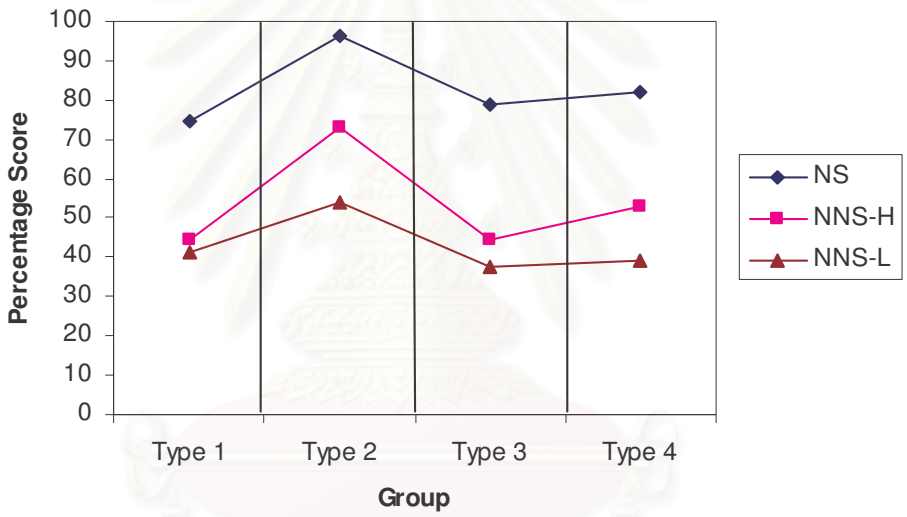
As there were 12 items for each NP type, with 30 participants in each group, the number of tokens for each type by group was 360, as shown in Table 5.3. The results indicated that there was a sizable gap in performance on each NP type between the NS and NNS-H groups. Of the two nonnative groups, the NNS-H students outscored the NNS-L group in every NP type. However, the gap in the percentage scores between the high and low-proficiency learners in Type 1 (generics) was relatively small (i.e. 44.2% and 41.4%). The low score of 44.2% obtained by the NNS-H group indicated that Type 1 was problematic for Thai learners even at a high proficiency level. Of the four NP types, Type 2 (referential definites) revealed the highest scores performed by all groups (as shown in bold). At first glance, it appears as if Type 2 was the least problematic. However, as this NP type is marked only with *the* while the other types are marked with at least two articles, we may also need to look more closely into the performance on each article in all NP types to see which article is more problematic than the others. This will be investigated in the next section.

Figure 5.2 below gives a clearer visual format of how each group performed on each NP type. One can see that the performance of the NS group varied across NP types, which suggests that the NSs found some NP types more complicated than others. It appears that the targetlike use of articles in the NS and NNS-H groups followed the same pattern. Both groups performed best on Type 2 (referential definites), followed by Type 4 (non-referentials), Type 3 (referential indefinites), and Type 1 (generics). In the NNS-L group, the ordering of NP types differed in that they produced targetlike uses for Type

1 (generic) articles at a higher rate than Type 4 and Type 3 articles. Thus, their pattern was the following (starting with the highest scores): Type 2 (referential definites), Type 1 (generics), Type 4 (non-referentials), and Type 3 (referential indefinites). The ordering of test performance on NP types in the three participant groups is shown below:

- NS: Types 2 > 4 > 3 > 1
- NNS-H: Types 2 > 4 > 3 > 1
- NNS-L: Types 2 > 1 > 4 > 3

Figure 5.2: Test Scores in Percentage by Group and NP Type



From Figure 5.2, one can observe that a substantial gap existed between the native speakers and Thai learners in each NP type. However, between the high- and low-proficiency learners, Type 2 displays a markedly larger gap than the other NP types, while the gaps in Type 1 and Type 3 were relatively small. In order to see whether or not such an observation is correct, an ANOVA was introduced. There were significant differences between groups for each NP type: Type 1, $F(2, 87) = 54.667$, $p < .001$; Type 2, $F(2, 87) = 68.598$, $p < .001$; Type 3, $F(2, 87) = 62.731$, $p < .001$; and Type 4, $F(2, 87) = 55.694$, $p < .001$. As *post hoc* tests, Tukey's HSD was introduced to examine significant differences in each NP type across groups. The results showed that there were differences in performance among groups, except between the NNS-H and NNS-L groups in Type 1 and Type 3, as shown in Table 5.4 below.

Table 5.4: *Comparisons of the Mean of Raw Scores between Groups by NP Type*

NP Type	Group Comparison	Difference of Means of Raw Scores	Sig.
Type 1	NS vs. NNS-H	3.63*	.000
	NS vs. NNS-L	3.97*	.000
	NNS-H vs. NNS-L	0.33	.709
Type 2	NS vs. NNS-H	2.80*	.000
	NS vs. NNS-L	5.10*	.000
	NNS-H vs. NNS-L	2.30*	.000
Type 3	NS vs. NNS-H	4.10*	.000
	NS vs. NNS-L	4.90*	.000
	NNS-H vs. NNS-L	0.8	.209
Type 4	NS vs. NNS-H	3.43*	.000
	NS vs. NNS-L	5.17*	.000
	NNS-H vs. NNS-L	1.73*	.002

* The mean difference is significant at the .05 level

5.1.3 Difference in Performance on the Article Test by NP Sub-Types

In order to examine the choice of each article that marks each NP type (hereafter referred to as *NP sub-type*), test scores were also calculated and are presented in Table 5.5. Figure 5.3 also shows the same set of scores in a line graph.

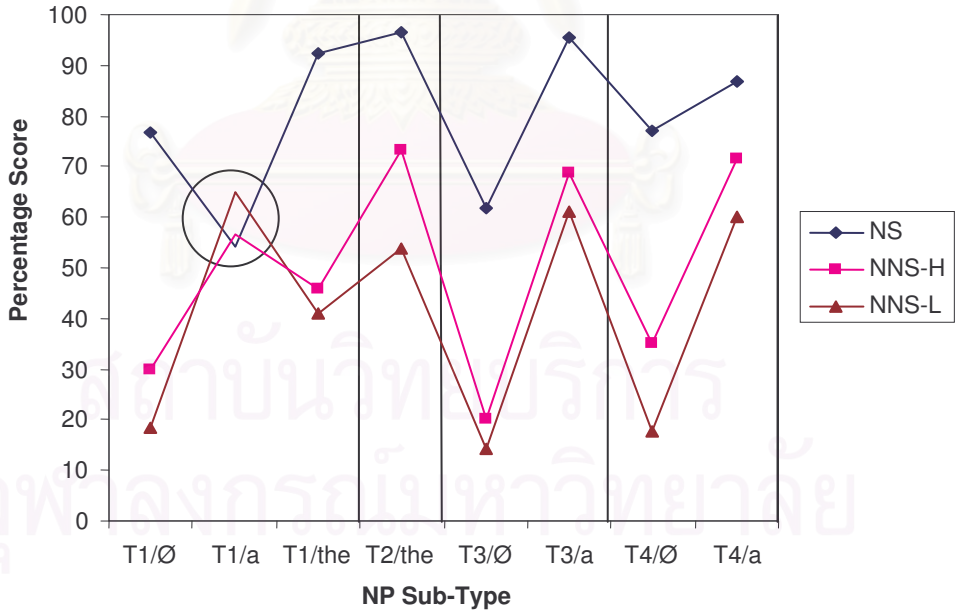
Table 5.5: *FD and Percentage of Target Article Use by Group and NP Sub-Type*

Group	NP Type							
	Type 1 (N=360)			Type 2 (N=360)	Type 3 (N=360)		Type 4 (N=360)	
	T1/Ø (N=120)	T1/a,an (N=120)	T1/the (N=120)	T2/the (N=360)	T3/Ø (N=180)	T3/a,an (N=180)	T4/Ø (N=180)	T4/a,an (N=180)
NS	92 (76.7%)	**65 (54.2%)**	111 (92.5%)	347 (96.4%)	111 (61.7%)	172 (95.6%)	139 (77.2%)	156 (86.7%)
NNS-H	36 (30.0%)	**68 (56.7%)**	55 (45.8%)	263 (73.1%)	36 (20.0%)	124 (68.9%)	63 (35.0%)	129 (71.7%)
NNS-L	22 (18.3%)	**78 (65.0%)**	49 (40.8%)	194 (53.9%)	26 (14.4%)	110 (61.1%)	32 (17.8%)	108 (60.0%)

As described in the above section, each NP type contained 360 tokens for each group of participants. With different articles marking each NP type, the number of tokens for each NP sub-type varied. For example, *Ø, a/an,* and *the* mark Type 1 NPs with 4 items representing each article; thus, the number of tokens for each article in

items in this NP type was 120. In Type 2, *the* marks 12 test items; thus, the number of tokens for *the* in items in Type 2 was 360. Type 3 and Type 4 are both marked with two articles: *Ø,* and *a/an*. Accordingly, the number of tokens for each article in items in Type 3 and Type 4 was 180. As Table 5.5 indicates, scores on the target use of article in the NS group varied greatly across the eight NP sub-types. Scores ranged from the lowest at 54.2% in T1/a,an to the highest at 96.4% in T2/the. A substantial gap in scores existed between the NS and the NNS-H groups in most NP sub-types (except in T1/a,an). Overall, the NNS-H group outperformed the NNS-L group in each NP sub-type. However, in T1/a,an (as displayed in bold), the pattern appeared to be reversed. The low-proficiency learners surprisingly produced the highest rate of the target article for this NP sub-type (i.e. 65.0%), and the high-proficiency learners outperformed the native speakers (i.e. 56.7% > 54.2%). This tends to suggest that generic-*a/an* was a fuzzy NP sub-type of generic articles.

Figure 5.3: Test Scores in Percentage by Group and NP Sub-Type



As shown in Figure 5.3, the targetlike use of articles in the NNS-H group followed a similar pattern to that of the native speakers across most NP sub-types, except T1/a,an in the circle. The pattern in the NNS-L group differed more from that of the NS group, particularly in their performance on T1/a,an in the circle. Type 3 NPs marked

with Ø (T3/Ø) were likely to yield the highest level of complication, at least among Thai learners. The scores of both NNS groups in T3/Ø were lowest among all NP sub-types. Also, the gap between the two groups was rather small. While the definite article use in T2/the yielded the highest scores in the performance of the NS and the NNS-H groups, this sub-type was ranked fourth in the targetlike use rate performed by the NNS-L group. The first three NP sub-types correctly produced by the NNS-L learners were those marked with the indefinite article *a/an* (i.e. in Type 1, Type 3 and Type 4).

In order to see the differences in scores among the eight NP sub-types on the article test, a one-way ANOVA was used to test the performances among the three sample groups. Details are illustrated in Table 5.6 below.

Table 5.6: *Comparisons of the Mean of Raw Scores between Groups by NP Sub-Type*

Sub-Type	Group Comparison	Difference of Means of Raw Scores	Sig.
T1/Ø	NS vs. NNS-H	1.87*	.000
	NS vs. NNS-L	2.33*	.000
	NNS-H vs. NNS-L	.47	.279
T1/a,an	NS vs. NNS-H	.10	.941
	NS vs. NNS-L	.43	.322
	NNS-H vs. NNS-L	.33	.509
T1/the	NS vs. NNS-H	1.87*	.000
	NS vs. NNS-L	2.07*	.000
	NNS-H vs. NNS-L	.20	.597
T2/the	NS vs. NNS-H	2.80*	.000
	NS vs. NNS-L	5.10*	.000
	NNS-H vs. NNS-L	2.30*	.000
T3/Ø	NS vs. NNS-H	2.50*	.000
	NS vs. NNS-L	2.83*	.000
	NNS-H vs. NNS-L	.33	.610
T3/a,an	NS vs. NNS-H	1.60*	.000
	NS vs. NNS-L	2.07*	.000
	NNS-H vs. NNS-L	.47	.312
T4/Ø	NS vs. NNS-H	2.53*	.000
	NS vs. NNS-L	3.57*	.000
	NNS-H vs. NNS-L	1.03*	.024
T4/a,an	NS vs. NNS-H	.90*	.020
	NS vs. NNS-L	1.60*	.000
	NNS-H vs. NNS-L	.70	.089

* The mean difference is significant at the .05 level

It was found that there were statistically significant differences in performance between groups in each sub-type except in Type 1, marked with *a/an* (T1/a,an): T1/Ø, $F(2,87) = 33.010$, $p < .001$; T1/a,an, $F(2,87) = 1.146$, $p < .3$, T1/the, $F(2,87) = 61.174$, $p < .001$; T2/the, $F(2,87) = 68.598$, $p < .001$; T3/Ø, $F(2,87) = 38.982$, $p < .001$; T3/a,an, $F(2,87) = 23.229$, $p < .001$; T4/Ø, $F(2,87) = 44.953$, $p < .001$; T4/a,an, $F(2,87) = 11.968$, $p < .001$. Tukey's HSD further revealed the following results (Please refer to Table 5.6).

(1) The scores of the three groups were significantly different in two NP sub-types: T2/the and T4/Ø. This means that the scores on T2/the and T4/Ø could discriminate the proficiency levels of all three participant groups.

(2) The difference was not significant among all groups in one NP sub-type: T1/a,an. This indicates that the scores of the three groups on T1/a,an could not discriminate between the native speakers and the two groups of Thai learners, nor could the scores discriminate the proficiency levels between the NNS-H and NNS-L groups.

(3) In the remaining five NP sub-types—i.e. T1/Ø, T1/the, T3/Ø, T3/a,an, and T4/a,an—the difference was statistically significant between the NS group and the two NNS groups, but not significant between the NNS-H and NNS-L groups. This means that the scores on these five NP sub-types could discriminate between the native speakers and Thai learners. However, the scores could not discriminate the proficiency levels between the NNS-H and the NNS-L groups.

In order to understand why the participants performed as they did, the quantitative results presented in this chapter will be further analyzed and discussed qualitatively in Chapter 6.

Up to this point, the participants' performance by each article that marks each NP type has been examined. From the results presented thus far, one could assume that the high percentage of the targetlike uses of *the* in the NNS-H group and *a/an* in the NNS-L group indicated that the NNS-H group had a better control of *the* than the other articles, while the NNS-L group had an earlier mastery of *a/an* than *the* and *Ø*. However, such an assumption is merely based on the learners' test performance. There is no guarantee that correct uses of articles among the Thai learners were based on their true understanding of the article system as researchers and teachers might expect. It could be the case that the

NNS-L learners over-generalized the indefinite article *a/an* in most NP contexts, due to the fact that *a/an* is normally introduced earlier in beginning EFL courses. Less proficient learners probably begin to use this article with all singular nouns. With increasing proficiency, learners may familiarize themselves more with using the definite article and may begin to over-generalize *the*. Such learners may feel more comfortable with using *the* as they recognize that *the* is neutral for countability and number. The learner's over-generalization of one form could be instrumental in boosting accuracy rates for contexts that require a certain article. Over-generalization of an article may accordingly be interpreted to mean that acquisition of that article was delayed. In order to understand the learners' performance better, Chapter 6 will investigate in more detail the metalinguistic explanations of the participants in the use of these articles in their given contexts.

In the following section, the participants' metalinguistic reasoning for their target article use on the fill-in-the-article test was investigated and analyzed quantitatively. The analyses in this section were conducted in order to explore the following questions:

(1) What is the metalinguistic knowledge (MLK) in English article use employed by native English speakers and Thai EFL learners of high and low proficiency levels?

(2) What are the similarities and differences in the MLK of English articles used by native English speakers and Thai EFL learners of high and low proficiency levels?

5.2 Difference in Metalinguistic Reasoning

5.2.1 Difference in Metalinguistic Reasoning for the Target Article Use

The quantitative analyses of metalinguistic reasoning in this section were conducted based on the classification of the participants' metalinguistic explanations as described in Chapter 4. The explanations were coded based solely on the actual statements that the participants made without the researcher making any assumptions on their behalf. These reasons were first classified as "explicit" (i.e. participants explicitly identified rules or other reasons for their article choices) or "non-explicit" (i.e. participants did not explicitly identify any specific reasons for their article choices). The first group consisted of reasons that were: (1) pragmatically-based (P), (2) semantico-

syntactically-based (S), and (3) the combination of both (P) and (S). It was found that when the participants mentioned both (P) and (S) together, they interchangeably stated (P) before (S) or (S) before (P) without a specific pattern. This accordingly suggested that such explanations should be coded as one sub-category, independent of (P) or (S). The assignment of (PS) for this sub-category does not indicate that (S) was a sub-type of (P) or vice versa. (Please refer to Section 4.5 in Chapter 4 for samples of participants' explanations in these three sub-categories of 'explicit' reasons)

The second group of explanations—i.e. the 'non-explicit' reasons—comprised statements based on the following three factors: (1) participants' intuitions or feelings (I), (2) guessing or pseudo-guessing (G), and (3) no reasons being given (N). (See Section 4.5 in Chapter 4 for examples of 'non-explicit' reasons in these three sub-categories)

Table 5.7 illustrates the distribution and percentage of metalinguistic reasoning for the target article use of the 48 test items by the three groups of participants.

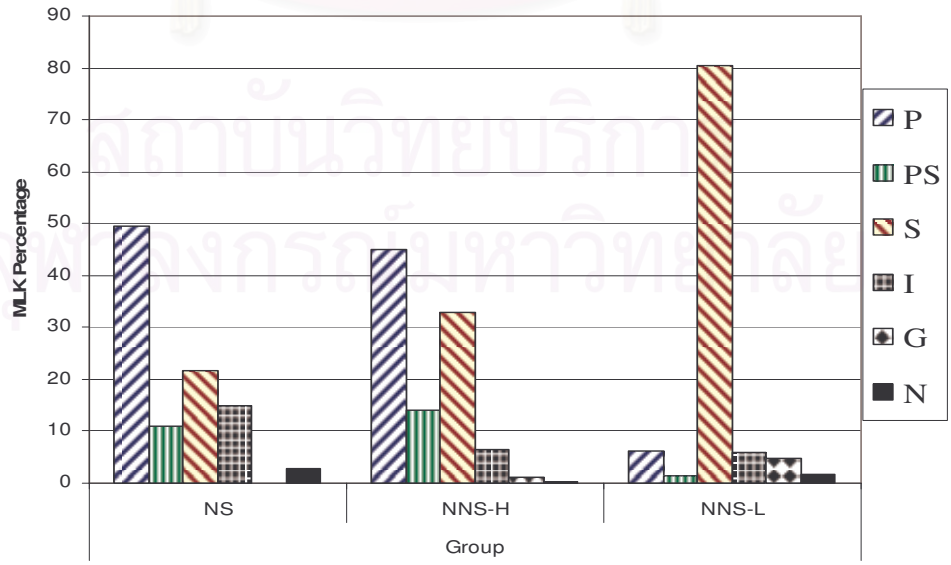
Table 5.7: *Metalinguistic Reasoning for Target Article Use*

MLK		Group		
		NS (N = 1,193/1,440)	NNS-H (N = 774/1,440)	NNS-L (N = 619/1,440)
Explicit	P	**591 (49.5%)**	**349 (45.1%)**	38 (6.1%)
	PS	131 (11.0%)	109 (14.1%)	8 (1.3%)
	S	258 (21.6%)	256 (33.0%)	**497 (80.3%)**
Total		980 (82.1%)	714 (92.2%)	543 (87.7%)
Non-Explicit	I	178 (14.9%)	50 (6.5%)	36 (5.8%)
	G	-	8 (1.0%)	30 (4.9%)
	N	35 (2.9%)	2 (0.3%)	10 (1.6%)
Total		213 (17.8%)	60 (7.8%)	76 (12.3%)
Grand Total		1,193 (100%)	774 (100%)	619 (100%)

As the test consisted of 48 items, with 30 participants in each group, the number of tokens for the total target use of articles in each group equaled 1,440. In the NS group, the frequency count for the MLK of the target article use was 1,193, whereas the MLK of the NNS-H group and the NNS-L group were 774 and 619 respectively. The data in Table 5.7 indicate that the NS group explained their target use of articles explicitly based on pragmatic (P) MLK at the highest percentage (49.5%), as shown in bold, followed by semantico-syntactic (S) MLK at 21.6%, and the combination (PS) at

11.0%. Non-explicit reasons given by the NS group constituted a high proportion (i.e. (I) = 14.9%, (N) = 2.9%, the total of 17.8%). The NNS-H students' use of MLK was similar to the NS group. They explained their article choices based on pragmatic (P) MLK at the highest percentage (45.1%), as shown in bold, followed by semantico-syntactic (S) MLK (33.0%), and the combination (PS) at 14.1%. The total percentage of non-explicit reasons, classified as (I), (G), and (N), supplied by the NNS-H students was 7.8%, which was lower than in the NS group. For the NNS-L learners, semantico-syntax (S) MLK was used at a substantially higher rate (i.e. 80.3%), as shown in bold, than all other MLK sub-categories. Unlike the NS and NNS-H groups, the NNS-L learners rarely used (P) and (PS) MLK to explain their article choices. The data showed that non-explicit reasons, classified as (I), (G), and (N) given by the NNS-L students came to a total of 12.3%. This suggests that the NNS-L group gave non-explicit explanations more often than the NNS-H students but with a lower frequency than the NS group. The results also indicated that the low-proficiency students selected articles based on guessing (G) at a higher percentage than the high-proficiency students (i.e. 4.9% > 1.0%). This point will be discussed in more detail in Chapter 6. The graph shown in Figure 5.4 below gives a clearer visual representation of the percentages of MLK sub-categories for the target use of articles across the groups.

Figure 5.4: *Metalinguistic Reasoning for Target Article Use by Group*



5.2.2 Difference in Metalinguistic Reasoning for the Target Article Use by NP Type

In order to examine how the participants justified their target article use in different NP contexts, descriptive statistics were used to calculate the percentages of metalinguistic reasoning provided by each group for each NP type, as displayed in Table 5.8.

Table 5.8: *Percentage of Metalinguistic Reasoning for Target Article Use by NP Type*

Native English Speakers (NS)					
MLK		NP Type			
		Type 1 (N=268)	Type 2 (N=347)	Type 3 (N=283)	Type 4 (N=295)
Explicit	P	**54.5**	**69.4**	31.5	**39.0**
	PS	11.2	4.6	19.4	10.2
	S	17.2	10.7	**32.5**	28.1
Non-Explicit	I	13.4	12.7	14.5	19.3
	G	-	-	-	-
	N	3.7	2.6	2.1	3.4
High-proficiency Learners (NNS-H)					
MLK		NP Type			
		Type 1 (N=159)	Type 2 (N=263)	Type 3 (N=160)	Type 4 (N=192)
Explicit	P	**51.6**	**62.0**	34.4	25.5
	PS	13.8	6.8	18.7	20.3
	S	30.2	23.2	**41.3**	**42.2**
Non-Explicit	I	4.4	6.5	4.4	9.9
	G	0	1.5	1.2	1.0
	N	0	-	-	1.0
Low-proficiency Learners (NNS-L)					
MLK		NP Type			
		Type 1 (N=149)	Type 2 (N=194)	Type 3 (N=136)	Type 4 (N=140)
Explicit	P	4.0	13.9	2.2	1.4
	PS	2.0	1.0	1.5	0.7
	S	**79.2**	**75.8**	**85.3**	**82.9**
Non-Explicit	I	6.7	3.1	6.6	7.9
	G	6.0	3.6	3.7	6.4
	N	2.0	2.6	0.7	0.7

According to Table 5.8, the total number of tokens for the target article use in each NP type equals 360 (30 participants responding to 12 test items). The metalinguistic reasoning is presented in a percentage calculated from the number of tokens of the target article use for each NP type by group as shown in parentheses. For example, the target article use in the NS group yielded 268 tokens in Type 1 (generics),

347 tokens in Type 2 (referential definites), 283 tokens in Type 3 (referential indefinites), and 295 tokens in Type 4 (non-referentials). As shown in Table 5.8, the native speakers explained their target article choices using pragmatic (P) MLK at a high percentage in each NP type. In Type 2 (referential definites), the pragmatic (P) MLK used by the NS group was highest (69.4%), followed by Type 1 generics (54.5%) and Type 4 non-referentials (39.0%), as displayed in bold. In Type 3 (referential indefinites), the use of semantico-syntactic (S) MLK and pragmatic (P) MLK was very close (i.e. 32.5% and 31.5%) although the (S) MLK was used at a slightly higher rate than the (P) MLK. It is notable that in Type 3, (PS) MLK was also used at a relatively high percentage (19.4%) by the native speakers. This suggests that Type 3 appeared to require semantic and/or syntactic justifications as well as pragmatic considerations in determining article choices. This is also the case with Type 4 NPs. As Table 5.8 indicates, although pragmatic (P) MLK was employed most frequently by the native speakers in Type 4 (i.e. in 39.0%), (S) and (PS) explanations were offered at substantial percentages (i.e. 28.1% and 10.2%) by the NS participants.

Non-explicit reasons, classified as (I) and (N), were given most frequently in Type 4 (non-referential) contexts (i.e. the total of 22.7%), followed by Type 1, Type 3 and Type 2 (i.e. 17.1%, 16.6%, and 15.3% respectively). That the native speakers had more difficulty in providing explicit explanations for their article choices than the Thai learners could result from the effect of article instruction which was less explicit in L1 classes than in EFL article lessons. As discussed earlier in Chapter 3, Type 4, Type 1, and Type 3 NPs are marked with *a/an* and *Ø*, which are articles that compile multiple functions (i.e. referentiality, countability and number). Thus, it may well be speculated that native speakers should have more difficulties articulating reasons for using these articles than rationalizing the use of *the* in Type 2 NPs. Since the use of *the* relies almost entirely on the discourse of communication while the distinctions of countability and number are neutralized, the native speakers tended to have less difficulties providing reasons for Type 2 article uses than other NP types. This will be discussed in greater detail in Chapter 6.

Among the Thai participants, the data in Table 5.8 indicate that the NNS-H students' use of MLK in explaining their article choices was similar to the NS's MLK. Pragmatic (P) MLK was used at a high proportion by the NS and NNS-H groups in each

NP type, with the highest percentage in Type 2, followed by Type 1. This suggests that the selective process of Type 2 (referential definite) and Type 1 (generic) articles depended heavily on contextual information. In Type 4 and Type 3, marked with *a/an* and Ø, semantico-syntactic (S) MLK was used most often in the NNS-H students' article choices (42.2% and 41.3%), followed by (P) MLK in 25.5% in Type 4 and 34.4% in Type 3, and (PS) MLK in 20.3% and 18.7%. As earlier discussed, *a/an* and Ø in Types 3 and 4 incorporate the notions of countability and number in addition to referentiality. Thus, it is not surprising that (S) and (PS) reasons were cited at a substantial rate among the NNS-H students in Type 3 and Type 4 contexts.

Non-explicit reasons, classified as (I), (G), and (N), were offered most frequently in the NNS-H group for Type 4 articles, followed by Type 2, Type 3, and Type 1. Overall, non-explicit reasons constituted lower rates in the NNS-H group than in the NS group. This indicated that the NNS-H students explained their article choices based on explicit reasons more often than the native speakers. This will be discussed in more detail in Chapter 6.

In the NNS-L group, the students explained their target article choices in each NP type based on semantico-syntactic (S) MLK at a remarkably high percentage (Please refer to Table 5.8). Pragmatically-based (P) MLK was used more frequently with Type 2 than the other NP types. As can be seen, explanations based on (P) and (PS) MLK were offered at very small percentages, particularly with Type 4 articles.

Non-explicit reasons (classified as I, G, and N) were supplied most often by the NNS-L group in Type 4, followed by Type 1, Type 3, and Type 2. Overall, the NNS-L supplied non-explicit reasons more frequently than the NNS-H group in each NP type. However, if we compare the total percentage of non-explicit reasons (i.e. I + G + N) offered for each NP type between the two NNS groups, we find that Type 1 demonstrated the highest rate of discrepancy. The NNS-L students gave non-explicit reasons for Type 1 article choices at 14.7%, while the NNS-H students were at only 4.4%. This suggests that although the NNS-L group chose the target articles for Type 1 items at the rate closer to that supplied by the NNS-H group, it does not mean that they had clear explanations for their article choices. When compared with the NS group, the NNS-L students gave non-explicit explanations less frequently than the NS group for

each NP type. This reflects perhaps the way in which article instruction is provided in Thai EFL settings—i.e. through explicit rule-based teaching. This will be discussed when reporting the qualitative analyses in Chapter 6.

In summary, the use of MLK among the NNS-H students was similar to the NS group. As Type 2 and Type 1 articles rely heavily upon contexts, the NS and NNS-H groups explained their choices of these two NP types based on pragmatic (P) MLK most frequently. Both groups used (S) MLK to justify their choices of Type 3 articles more often than other types of MLK. The MLK reasoning for Type 4 context differed to some extent between the two groups. The native speakers explained their choice with (P) MLK more often than (S), whereas the NNS-H group used (S) MLK to explain their Type 4 article choices more frequently than (P). In the NNS-L group, the students' use of MLK was different. They explained their article choices in all NP types based on (S) MLK at the highest rate. Other types of explicit MLK were used in very small percentages by the NNS-L students. Of the three groups, the NS participants gave non-explicit reasons most frequently, whereas the NNS-H students offered non-explicit explanations the least. Type 4 articles constituted the highest proportion of non-explicit reasons cited by all groups. As Type 4 articles compile several complicated functions (i.e. referentiality, countability, and number), it may be hypothesized that a large number of participants had difficulty explaining why the articles they chose should be appropriate in the given contexts.

5.2.3 Difference in Metalinguistic Reasoning for the Target Article Use by NP Sub-Type

In order to examine what kind of metalinguistic knowledge the participants used in determining the article that marks each NP type (referred to herein as *NP sub-type*), the quantitative analyses were conducted to provide the percentages of metalinguistic reasoning for the target article choices by each group of participants. The results revealed the similarities and differences in the MLK used between the native speakers and Thai learners with high and low proficiency levels, as displayed in Table 5.9 below.

Table 5.9: *Percentage of Metalinguistic Reasoning for Target Article Use by NP Sub-Type*

Native English Speakers (NS)									
MLK		NP Sub-Type							
		T1/Ø N=92	T1/a,an N=65	T1/the N=111	T2/the N=347	T3/Ø N=111	T3/a,an N=172	T4/Ø N=139	T4/a,an N=156
Explicit	P	**43.5**	**41.5**	**71.2**	**69.4**	**37.8**	27.3	**42.5**	35.9
	PS	7.6	18.5	9.9	4.6	8.1	26.7	5.0	14.7
	S	17.4	27.7	10.8	10.7	23.4	**38.4**	18.7	**36.5**
Non-Explicit	I	23.9	9.2	7.2	12.7	25.2	7.6	27.3	12.2
	G	-	-	-	-	-	-	-	-
	N	7.6	3.1	0.9	2.6	5.4	-	6.5	0.6
High-Proficiency Learners (NNS-H)									
MLK		NP Sub-Type							
		T1/Ø N=36	T1/a,an N=68	T1/the N=55	T2/the N=263	T3/Ø N=36	T3/a,an N=124	T4/Ø N=63	T4/a,an N=129
Explicit	P	16.7	35.3	**94.5**	**62.0**	2.8	**43.5**	9.5	33.3
	PS	27.8	17.6	-	6.8	22.2	17.7	20.6	20.2
	S	**38.9**	**45.6**	5.5	23.2	**61.1**	35.5	**52.4**	**37.2**
Non-Explicit	I	16.7	1.5	-	6.5	8.3	3.2	15.9	7.0
	G	-	-	-	1.5	5.6	-	-	1.6
	N	-	-	-	-	-	-	1.6	0.8
Low-Proficiency Learners (NNS-L)									
MLK		NP Sub-Type							
		T1/Ø N=22	T1/a,an N=78	T1/the N=49	T2/the N=194	T3/Ø N=26	T3/a,an N=110	T4/Ø N=32	T4/a,an N=108
Explicit	P	-	1.3	10.2	13.9	-	2.7	-	1.9
	PS	-	-	6.1	1.0	-	1.8	-	0.9
	S	**54.5**	**94.9**	**65.3**	**75.8**	**65.4**	**90.0**	**56.3**	**90.7**
Non-Explicit	I	22.7	-	10.2	3.1	15.4	4.5	31.1	0.9
	G	18.2	3.8	4.1	3.6	19.2	-	12.5	4.6
	N	4.5	-	4.1	2.6	-	0.9	-	0.9

As earlier described in Section 5.1.3, the number of tokens for the target use of article in T1/Ø, T1/a,an, and T1/the was 120 for each sub-type; the number of tokens for T2/the was 360; and the number of tokens for each of T3/Ø, T3/a,an, T4/Ø, and T4/a,an was 180. The percentage of metalinguistic reasoning was calculated from the number of tokens of the target article use for each NP sub-type by group. The highest percentage for each NP sub-type is displayed in bold type.

The data in Table 5.9 indicate that the NS group used pragmatic (P) MLK most frequently in explaining their targetlike article uses in almost all NP sub-types, with the highest percentage in T1/the at 71.2% followed by T2/the at 69.4%. As reviewed in the literature, the use of *the* in definite descriptions and generic uses require pragmatic

considerations of the contexts in making article choices. This is supported by the results of the MLK used by the native speakers in this study, whereby their article uses in Type 2 and Type 1 items were explained based on pragmatic (P) MLK in the highest percentage. With regard to the target article *Ø*, pragmatic (P) MLK comprised the highest proportion in each NP type in the NS group (i.e. 43.5%, 37.8%, and 42.5% in Type 1, Type 3 and Type 4 respectively), with (S) MLK constituting the second highest percentage (i.e. 17.4%, 23.4%, and 18.7%). For the targetlike use of *a/an*, semantico-syntactic (S) MLK yielded the highest percentage in Type 3 and Type 4 (38.4% and 36.5%). (PS) MLK was also used frequently by the NS group in justifying their choice of *a/an* in all contexts. As the distinction between the use of *a* and *an* requires phono-syntactic consideration, which was categorized as (S) MLK, it was found that the NS choice of *an* was explained based mainly on this phono-syntactic rule. This seems to boost the percentage of using (S) and (PS) MLK in the choice of *a/an* in the NS group. However, although the proper use of *a/an* (as well as *Ø*) requires considerations of other important features such as countability and number, one can see that (P) MLK constituted a high proportion in the native speakers' choice of *a/an*, not to mention the highest rate of using (P) MLK to explain their choice of *Ø*. This suggests that even though countability and number are involved in the use of *a/an* and *Ø*, the native speakers still regarded the pragmatic aspect as a crucial feature in determining articles.

As I mentioned in the previous section, the native speakers gave non-explicit explanations (I + N) more frequently than the Thai participants. From Table 5.9 above, one can see that the percentage of non-explicit reasons was substantial in the native speakers' choice of *Ø*. Non-explicit explanations were highest in the NS use of *Ø* in Type 4 (I = 27.3%, N = 6.5%, the total of 33.8%), followed by Type 1 (I = 23.9%, N = 7.6%, the total of 31.5%), and Type 3 (I = 25.2%, N = 5.4%, the total of 30.6%). As related to this, the MLK explanations were analyzed qualitatively and will be discussed in detail in Chapter 6.

In summary, the native speakers used (P) MLK more frequently than other MLK sub-categories when they chose *the* and *Ø*. With their choice of *Ø*, (S) and (PS) MLK was also used substantially. The NS participants cited (S) and (PS) reasons most frequently with their choice of the indefinite article *a/an*, but also with high percentages

of (P) MLK. The *zero* article (*Ø*) was the article for which non-explicit explanations were supplied the most.

Among the Thai learners, the MLK used in explaining article choices in the NNS-H students was more consistent with the MLK used by the native speakers than in the NNS-L group. (P) MLK was used most among the NNS-H students in their targetlike choice of *the* in Type 1 and Type 2 contexts (94.5% and 62.0%). With *a/an* and *Ø* in Type 3, Type 4, and Type 1 (which required the detection of countability and number), the NNS-H students used (S) MLK more frequently to support their choices. (PS) explanations were also cited in high proportions when the students used *a/an* and *Ø* in these NP types. The use of MLK among these students could result from the effect of EFL article instruction whereby the focus is often on countability and number when discussing nouns and noun phrases in relation to article usage. Interestingly, however, it should be noted that in Type 3 marked with *a/an*, (P) MLK was used to explain the students' choice at a higher percentage than (S) MLK (43.5% > 35.5%), with the combination (PS) MLK at 17.7%. It could be hypothesized that first-mention (Type 3) NPs, which are normally used in a narrative text, were likely to direct the students' attention toward referentiality rather than merely the semantic or syntactic aspect. Despite this fact, (S) MLK was used more frequently when the students used *a/an* than when they used *the*.

Like the NS group, the NNS-H students' choice of *Ø* was explained based on non-explicit reasons more frequently than their choices of the other articles. The pattern was also the same—that is, the use of *Ø* in Type 4 yielded the highest percentage of non-explicit reasons given by the NNS-H students (I = 15.9%, N = 1.6%, the total of 17.5%), followed by Type 1 (I = 16.7%), and Type 3 (I = 8.3%, G = 5.6%, the total of 13.9%).

To sum up, consistent with the NS group, the NNS-H students used (P) MLK more frequently with their choice of the definite article *the*. They used (S) and (PS) with their choices of the indefinite article *a/an* and *Ø*. It was only with the use of *a/an* in Type 3 that the NNS-H students used (P) MLK more often than (S). For the target use of *Ø*, similar to the NS group, non-explicit explanations were found with the use of this article more frequently than with the other articles.

For the NNS-L learners, their MLK explanations for the target article choices were different from the MLK used in the NS and NNS-H groups. The NNS-L students explained their choices based primarily on semantico-syntactic (S) MLK in all NP sub-types (Please refer to Table 5.9). Even in Type 1 and Type 2 marked with *the*, whereby (P) MLK constituted the highest proportion in the explanations provided by the NS and NNS-H groups, the NNS-L students explained their choice using semantico-syntactic (S) MLK at substantially higher percentages than (P) MLK. Of the three articles, (S) explanations were used most frequently when the NNS-L students chose the indefinite article *a/an*, with the highest percentage found in T1/a,an (94.9%), followed by T4/a,an (90.7%) and T3/a,an (90.0%). (S) MLK was used with the students' choice of *the* at a substantially lower percentage (i.e. 75.8% in Type 2 and 65.3% in Type 1). As the low-proficiency students supplied a great number of non-explicit reasons for the choice of *Ø*, the use of (S) MLK with this article choice was lower than with the other articles—i.e. 54.5% in Type 1, 56.3% in Type 4, and 65.4% in Type 3. Overall, the NNS-L students used (P) and (PS) MLK in a small percentage in each NP sub-type. (P) MLK was used most with the students' choice of *the* in Type 2 at 13.9%, followed by the use of *the* in Type 1 at 10.2%. It may be noted that with the choice of *Ø* (in Type 1, Type 3, and Type 4), the NNS-L students did not use (P) or (PS) MLK at all.

Similar to the NNS-H and NS groups, the NNS-L students also gave non-explicit reasons with their targetlike choice of *Ø* at a high proportion in each NP type. The use of *Ø* in Type 1 (generics) constituted the highest portion of non-explicit reasons given by the NNS-L students (I = 22.7%, G = 18.2%, N = 4.5%, the total of 45.4%), followed by Type 4 (I = 31.1%, G = 12.5, the total of 43.6%), and Type 3 (I = 15.4%, G = 19.2%, the total of 34.6%). Of all the three articles, non-explicit reasons given for the use of *a/an* in the NNS-L group comprised the smallest percentage. With the targetlike use of *a/an*, non-explicit reasons (I + G + N) came to a total of 6.4% in Type 4, 5.4% in Type 3, and only 3.8% in Type 1.

5.3 Summary

This chapter presented the results of the quantitative analyses of the participants' performances on the article test and their metalinguistic explanations for their article choices. As stated earlier, this chapter aimed to examine the metalinguistic knowledge employed by native English speakers as opposed to the high- and low-proficiency Thai learners in using English articles in given contexts and to see whether or not there were differences in the learners' metalinguistic usage depending on English proficiency, using the native speakers' MLK as a baseline.

The analyses presented revealed that the scores on the test increased linearly according to the proficiency groups in each NP type. The test scores discriminated the performances between the NS and the two NNS groups in every NP type. However, the results showed that Type 1 (generic) NPs appeared to be the most problematic NP type, at least among the NNS-H students, as no significant difference was found in the performances between the high- and low-proficiency groups. Type 3 (referential indefinite) NPs also revealed no significant difference between the NNS-H and NNS-L groups. That the test scores did not discriminate the performances between the high- and low-proficiency students in Type 1 and Type 3 indicates that these two NP types are more complicated than the NP contexts in Type 2 and Type 4, at least among the NNS-H students. From the test results, all groups performed best on Type 2 items with the target article *the*.

When examining in greater depth the targetlike use of articles that mark each NP type, it was found that the test scores discriminated the participants' proficiency in all NP sub-types, except in T1/a,an. The statistical analysis revealed that the difference in the mean of raw scores in T1/a,an among the three groups was not significant. The test scores were found to decrease linearly with increasing proficiency levels. The only two sub-types whereby the scores discriminated the three proficiency groups were T2/the and T4/Ø. Between the two NNS groups, the results showed that there were no significant differences in test performance in T1/Ø, T1/the, T3/Ø, T3/a,an, and T4/a,an. As the scores of the NNS-H and NNS-L groups in these sub-types did not discriminate the proficiency levels, it indicated that these five sub-types were problematic and required more in-depth investigation into the MLK explanations provided by the Thai students.

In examining MLK used by the participants for their target article choices, the results showed that in terms of the overall performance on the fill-in-the-article test, the NNS-H students' use of MLK was more consistent with the native speakers' norms than when compared with the NNS-L students. Pragmatic (P) MLK was used most frequently in the target article choices of the NS and NNS-H groups. The NNS-L learners, unlike the other two groups, relied almost entirely on (S) MLK in explaining their target article choices on the test. They used pragmatic (P) MLK in a very low percentage.

With regard to NP types, the NS group used (P) MLK most frequently in three NP types: Type 2, Type 1, and Type 4 (with Type 2 yielding the highest percentage), while (S) MLK was used at a slightly higher rate than (P) MLK with Type 3 (first-mention) articles. In the NNS-H group, Type 2 and Type 1 article choices were explained based on pragmatic (P) MLK most often, whereas (S) MLK was used most frequently to explain their Type 3 and Type 4 target article choices. In the NNS-L group, (S) MLK was used at a substantially higher rate than the other MLK sub-categories in every NP type. When analyzing the MLK used for the target article choices in NP sub-types, the results showed that except for the choices of *a/an* in T3/a,an and T4/a,an, (P) MLK was used by the NS group in all NP sub-types. In T3/a,an and T4/a,an, although the NS group used (S) more frequently than (P), it was found that (P) and (PS) MLK also constituted a high proportion of the explanations of their targetlike choices. In the NNS-H group, (P) MLK was used most often with their targetlike choice of *the* in Type 1 and Type 2. The NNS-H students' choices of *a/an* and *Ø* were explained based primarily on (S) MLK, except with *a/an* in Type 3, whereby their target article choices were explained based on (P) MLK more often than (S). The use of MLK was different between the NS and NNS-H groups in the choice of *Ø*. The NNS-H group explained their *Ø* use based on (S) MLK at a substantially higher rate than (P), whereas the NS group used (P) MLK more frequently than (S) when they chose *Ø*. In the NNS-L group, the students explained their target article choices based on (S) MLK at a far higher percentage than the other MLK sub-categories in all NP sub-types.

The results indicate that the use of MLK between the NS and the two NNS groups was similar in that they tended to employ (S) MLK substantially when they chose *a/an*. In the choice of *the*, the frequent use of (P) MLK was similar in the NS and NNS-H groups, but different in the NNS-L group. In the choice of *Ø*, the use of (S) MLK was

similar in the NNS-H and NNS-L groups, but different in the NS group. Overall, the use of MLK in the NNS-H group was more consistent with the use of MLK in the NS group, particularly in the choice of the definite article *the*, which requires pragmatic (P) MLK for its selection.

In exploring non-explicit explanations supplied by the three groups, the results revealed that the NS group gave more non-explicit reasons than the Thai students of both proficiency levels. From the quantitative analyses, non-explicit reasons, particularly those based on impressionistic judgments (I), were used most frequently in the native speakers' choice of *Ø* in all contexts. Among the Thai students, the use of *Ø* also constituted the highest percentage of non-explicit reasons in both proficiency groups. Overall, the NNS-L group supplied non-explicit reasons for their article choices in a greater proportion than the NNS-H students. This suggests that the more proficient the students are, the more metalinguistic awareness in article usage they seem to have.

In this chapter, the quantitative analyses conducted on the participants' test performance and MLK used in explaining their article choices partly answer the two research questions cited earlier. The results support the corresponding hypotheses that the native speakers' MLK in article usage is primarily pragmatically-oriented. The MLK of the Thai students seems to be more syntactically- and semantically-based. However, it is clear that the MLK of the NNS-H students is more pragmatically-oriented than the MLK of the NNS-L students. As one can see, the NNS-H students explained their article choices of the definite article *the* based on (P) MLK at a high proportion. Even with their choice of *a/an*, (P) MLK and (PS) MLK were also used frequently by the NNS-H students to support this article choice. The use of MLK for *Ø* between the NS and the two NNS groups was different. The NNS-H and NNS-L students used (S) MLK most often when they chose *Ø*, while the NS group used (P) MLK at a higher rate than (S). The NNS-L students' MLK with regard to English articles was different from the NS and the NNS-H groups; it was based almost entirely on semantico-syntactic (S) orientations.

So far, the results have been presented based on quantitative analyses of the participants' performance on the test and the metalinguistic knowledge of the target article use. From the test scores, one cannot fully claim that EFL learners chose the target article based on linguistically appropriate reason(s). On the one hand, there is no

guarantee that learners have a proper understanding of article usage when they use articles correctly, while on the other hand, learners' inaccurate use of articles may not necessarily be due to their complete failure to understand the use of articles. Therefore, in order to capture the participants' true understanding of and performance with the English article system, in the next chapter, I will further present the results of the MLK explanations provided on the MLK elicitation questionnaire. I will focus mainly on the NP types and sub-types which were identified as problematic by the descriptive statistics as earlier presented (Please refer to Table 5.6 on page 97). In so doing, my goal is to understand the mental representations formulated and used in determining English articles between the native speakers and Thai learners of varying proficiency levels as well as to uncover the reasons why Thai EFL learners have so much difficulty in acquiring the English article system.

CHAPTER VI
RESULTS (2)
METALINGUISTIC EXPLANATIONS IN ENGLISH ARTICLE USE

The present chapter aims for a more in-depth examination of the metalinguistic knowledge (MLK) in the participants' article usage. The purpose of the analyses in this chapter is to further explore and discuss qualitatively the following two questions that were quantitatively analyzed and presented in the previous chapter.

(1) What is the metalinguistic knowledge (MLK) in English article use employed by native English speakers and Thai EFL learners of high and low proficiency levels?

(2) What are the similarities and differences in the MLK of English articles used by native English speakers and Thai EFL learners of high and low proficiency levels?

Based on the qualitative analyses conducted to investigate the above questions, the main goal of this chapter is to answer the final question addressed by this study:

(3) What are the problems underlying English article use in Thai learners, with regard to the MLK of definiteness, genericity, and countability?

The organization of the present chapter is as follows: first, I will provide an overview of how targetlike and non-targetlike article choices for all NP types were distributed across the groups. Then, I will further explore how variable answers were distributed among NP types and sub-types. Based on the participants' article choices, I will investigate the metalinguistic explanations provided by the NS and NNS learners for their article usage in order to understand why each group chose a certain article in its given context and to identify problems in article usage among Thai learners.

6.1 Article Uses on the Fill-in-the-Article Test

6.1.1 Article Uses in all NP Types

Table 6.1 below shows how article choices were distributed by test item representing each NP type and sub-type across the groups.

Table 6.1: *Distribution of Article Uses by Item in Each NP Type*

NP Type	Target Article	Test Item	NS Responses (N = 1,440)			NNS-H Responses (N = 1,440)			NNS-L Responses (N = 1,440)		
			Ø	a/an	the	Ø	a/an	the	Ø	a/an	the
	Ø	1	**22**	1	7	**5**	17	8		27	3
		11	**28**		2	**14**	2	14	**10**	13	7
		13	**21**		9	**12**		18	**8**	10	12
		39	**21**		9	**5**	1	24	**4**	8	18
	a/an	16		**20**	10		**18**	12		**22**	8
Type 1		40		**13**	17	1	**22**	7	1	**22**	7
		41		**15**	15	4	**15**	11	6	**9**	15
		42		**17**	13	2	**13**	15		**25**	5
	the	3		3	**27**		26	**4**		29	**1**
		4			**30**	2	10	**18**		16	**14**
		37		3	**27**	2	21	**7**	1	14	**15**
		45		3	**27**	1	**3**	**26**	3	8	**19**
	the	2			**30**		2	**28**		5	**25**
		8		1	**29**	7	15	**8**	6	14	**10**
		19			**30**	1	2	**27**	6	8	**16**
		22			**30**	5	4	**21**	9	14	**7**
		24			**30**	2	1	**27**	2	9	**19**
Type 2		27			**30**			**30**	2	5	**23**
		31	3		**27**	11		**19**	8	10	**12**
		35			**30**	1	7	**22**	1	7	**22**
		36			**30**	2	3	**25**	2	13	**15**
		38	8		**22**	5	3	**22**	2	16	**12**
		44			**30**	2	4	**24**	3	7	**20**
		48	1		**29**	6	14	**10**	2	15	**13**
	Ø	18	**25**	2	3	**2**	10	18	**2**	21	7
		21	**10**		20	**6**	3	21	**3**	1	26
		25	**14**		16		19	11	**2**	19	9
		29	**30**			**18**	7	5	**7**	19	4
		33	**17**	11	2	**5**	9	16	**4**	13	13
Type 3		34	**15**	14	1	**5**	6	19	**8**	15	7
	a/an	20		**30**		3	**19**	8	3	**18**	9
		23	1	**29**		2	**25**	3	3	**22**	5
		26		**27**	3	4	**22**	4	1	**21**	8
		28		**30**		6	**19**	5	3	**9**	18
		30		**29**	1	2	**18**	10		**18**	12
		32		**27**	3		**21**	9		**22**	8
	Ø	6	**20**		10	**11**	3	16	**7**	9	14
		7	**14**		16	**6**	4	20	**6**	14	10
		10	**22**	1	7	**6**		24	**3**	1	26
		12	**29**		1	**24**	1	5	**7**	7	16
		14	**27**		3	**3**	3	24	**5**	9	16
Type 4		15	**27**	1	2	**13**	10	7	**4**	21	5
	a/an	5		**30**		4	**25**	1	6	**23**	1
		9		**16**	14	3	**22**	5	4	**13**	13
		17	1	**27**	2	9	**16**	5	4	**13**	13
		43		**28**	2	3	**17**	10	3	**18**	9
		46		**25**	5	1	**26**	3	1	**23**	6
		47		**30**		3	**23**	4	8	**18**	4

As displayed in Table 6.1, the total number of tokens is 4,320—i.e. with 90 participants (30 in each group) responding to 48 items. Thus, the number of responses in each participant group is 1,440. Responses on the target article are marked in bold. The targetlike response of 30 for each test item indicates that all 30 students in that particular group chose the target article. Thus, for example, in Type 2 (referential definite) NPs, the targetlike use of *the* in the NS group yielded 30 tokens (i.e. 100% targetlike use) in almost all items, except items 8, 31, 38, and 48. The NS participants variably chose *a/an* in 1 token in item 8, while *Ø* was used in 1 token in item 48, 3 tokens in items 31, and 8 tokens in item 38. As the frequency of variable choices in the NS group was high in item 38, the performance on this item will be discussed later in Section 6.2.1.1. The qualitative analyses conducted for the participants' targetlike and non-targetlike article uses in this chapter aim to describe the performance of each group in a holistic fashion (as opposed to an item by item analysis). Therefore, generalizations will be made in terms of group performance and only notable points will be discussed.

To present an overview of how articles were distributed for all NP types among the three groups of participants, Table 6.2 below reorganizes the data in Table 6.1 and shows the percentages of the three articles: *Ø, a/an* and *the* used in all NP types across the groups.

Table 6.2: *Distribution of Article Uses in All NP Types by Percentage*

Target Article* (N = 1,440)	NS Article Choices (%)			NNS-H Article Choices (%)			NNS-L Article Choices (%)			Average Target Article (%)
	Ø	a/an	the	Ø	a/an	the	Ø	a/an	the	
Ø (480)	**71.2**	6.3	22.5	**28.1**	19.8	52.1	**16.7**	43.1	40.2	38.7
a/an (480)	0.4	**81.9**	17.7	9.8	**66.9**	23.3	8.9	**61.7**	29.4	70.2
the (480)	2.5	2.1	**95.4**	9.8	23.9	**66.3**	9.8	39.6	**50.6**	70.8

*Note: N for each article in each sample group is 480; Total N is 1,440.

The target use of each article by each group is shown in bold in Table 6.2. Thus, for example, for those items where *Ø* was the target form, the NS group chose *Ø* in 71.2%, the NNS-H in 28.1% and the NNS-L in 16.7%. The variable choice of *a/an* in the target *Ø* contexts was 6.3% in the NS, 19.8% in the NNS-H and 43.1% in the NNS-L, whereas the variable choice of *the* in the *Ø* contexts was 22.5%, 52.1% and 40.2% in

the NS, NNS-H, and NNS-L groups respectively. The average percentage in the targetlike choices for all articles comes to 82.8% in the NS group, 53.8% in the NNS-H group, and 43% in the NNS-L group, as earlier demonstrated in Table 5.1 on page 91 in Chapter 5. Generally speaking, the NS and NNS-H learners overused the definite article *the* (used *the* more than it was meant to be used), particularly in contexts where the *zero* article (*Ø*) was required. The indefinite article *a/an* was overused to a larger extent by the NNS-L learners in the target *Ø* and *the* contexts.

As one can see, *Ø* was used in the lowest rate across articles in all groups, particularly among the Thai learners. There was a substantial gap in the targetlike use of *Ø* between the NS and NNS groups. While the NS group produced 71.2 % targetlike use of *Ø*, the NNS-H and the NNS-L learners supplied only 28.1% and 16.7% respectively. The definite article *the* was chosen instead of *Ø* at a high proportion by all groups. This included almost a quarter of the responses supplied by the NS students (22.5%), about half of the responses produced by the NNS-H learners (52.1%), and 40% of the responses given by the NNS-L learners. The choice of *a/an* in the *Ø* contexts also constituted a high portion among Thai learners, particularly in the NNS-L group.

The average percentage of the targetlike use of each article in all groups was displayed in the last column of Table 6.2. As one can see, the participants used the target article *Ø* in the lowest proportion (38.7%). This lends evidence to the observation that *Ø* is the most problematic. The percentages of the targetlike uses of *the* and *a/an* by all groups were almost equal (70.8% and 70.2%). This is because the native speakers supplied the target article *the* at a higher percentage than *a/an* (i.e. 95.4% > 81.9%) while the two NNS groups supplied the target article *a/an* at a higher percentage than the target article *the* (i.e. 66.9% > 66.3% in the NNS-H group, and 61.7% > 50.6% in the NNS-L group).

In order to see how each article was used and for what reasons, we will examine the participants' article choices in definite, indefinite, and generic contexts based on the four NP types in the following section.

6.1.2 Article Uses in each NP Type

Table 6.3: *Distribution of Article Uses by NP Type by Percentage*

NP Type	Target Article (N=1,440)	NS Article Choices (%)			NNS-H Article Choices (%)			NNS-L Article Choices (%)		
		Ø	a/an	the	Ø	a/an	the	Ø	a/an	the
Type 1 (N=360)	Ø (120)	**76.7**	0.8	22.5	**30.0**	16.7	53.3	**18.3**	48.3	33.3
	a/an (120)	0	**54.2**	45.8	5.8	**56.7**	37.5	5.8	**65.0**	29.2
	the (120)	0	7.5	**92.5**	4.2	50.0	**45.8**	3.3	55.8	**40.8**
Type 2 (N=360)	the (360)	3.3	0.3	**96.4**	11.7	15.3	**73.1**	11.9	34.2	**53.9**
Type 3 (N=360)	Ø (180)	**61.7**	15.0	23.3	**20.0**	30.0	50.0	**14.4**	48.9	36.7
	a/an (180)	0.5	**95.6**	3.9	9.4	**68.9**	21.7	5.6	**61.1**	33.3
Type 4 (N=360)	Ø (180)	**77.2**	1.1	21.7	**35.0**	11.7	53.3	**17.8**	33.9	48.3
	a/an (180)	0.5	**86.7**	12.8	12.8	**71.7**	15.6	14.4	**60.0**	25.6

Table 6.3 demonstrates a consolidated picture of the participants' targetlike and non-targetlike uses of articles in each NP type in percentage. The targetlike uses of articles are displayed in bold type. In discussing the results of each group, I will report on the metalinguistic explanations in relation to the participants' targetlike and non-targetlike article uses for each target article in different NP contexts (i.e. definite, indefinite and generic contexts) based on Huebner's NP types. I will begin with the definite article *the*, which marks Type 2 NPs, followed by the indefinite articles *a/an* and *Ø*, which mark Type 3 (referential indefinite) and Type 4 (non-referential) contexts. Type 1 (generics), marked with all of the three articles *the*, *a/an* and *Ø*, appeared to be the most problematic NP type. Therefore, I will present generic article uses separately in the last section.

6.2 Metalinguistic Explanations in English Article Use

6.2.1 Metalinguistic Explanations in the Use of *the* in Definite Contexts

For definite NP contexts, the definite article *the* marks Type 2 [+SR, +HK] NPs whose referent is specific to the speaker [+SR] and assumed as known to (or identifiable by) the hearer [+HK] (Please refer to Chapter 3). Table 6.4 below displays the number of tokens of the targetlike and non-targetlike article uses in Type 2 (referential definite) contexts, with percentages in parentheses.

Table 6.4: *Frequency Distribution and Percentage of Article Uses in Type 2*

Type 2 [+SR +HK]				
Target Article		the (360 tokens)		
Article Choice		Ø	a/an	the
Group	NS	12 (3.3%)	1 (0.3%)	**347 (96.4%)**
	NNS-H	42 (11.7%)	55 (15.3%)	**263 (73.1%)**
	NNS-L	43 (11.9%)	123 (34.2%)	**194 (53.9%)**

As the purpose of the study is to use article choices and metalinguistic explanations of the native English speakers as a baseline to investigate Thai learners' performance and metalinguistic knowledge (MLK) used when making article choices, I will first present the test results and MLK of the NS group. Then, I will compare these results with those of Thai learners in the high and low English proficiency groups.

6.2.1.1 Metalinguistic Explanations of the Native Speakers in the Use of *the* in Definite Contexts

As displayed in Table 6.3 in Section 6.1.2, the NS group supplied the target article *the* in Type 2 contexts at the highest percentage among all NP types. The targetlike use of *the* reached 96.4%, while the variable uses of *a/an* and *Ø* constituted relatively small percentages. I will now report on the metalinguistic explanations given by the NS participants in their choices of the target article *the* and their variable choices of *a/an* and *Ø* in the context of Type 2 (referential definite) NPs.

(1) Use of the target article *the* in definite (Type 2) contexts

A very common explanation given by the NS group when they used *the* in Type 2 context indicated that most NS students interpreted given references as 'specific'. In their judgment of 'specific reference', the following factors were considered:

1.1 the existence of a modifier

1.2 a previously-mentioned NP

1.3 the association of a given referent with a previously-mentioned NP

In justifying the use of *the*, a number of native speakers considered the existence of modifiers as a device that marked definiteness. Thus, adjectives such as *best, same* and *first* attached to the given NPs were regarded as an indicator of 'specificity'. This can be seen in the following examples of comments made by three NS students.

(Quote 1) the (best thinker) – 'best' makes it a definite statement stating superiority

(Quote 2) the (same way) specifies which way is being spoken of

(Quote 3) the (first three months) identifies the specific time frame

The second factor that marked 'specificity' was the second or subsequent mention of a given NP in the same text. The following examples illustrate that the NS students who used *the* in front of previously-mentioned nouns such as *fire* and *hotel* based their choice on this reason.

(Quote 4) the (fire) - refers back to a specific fire previously mentioned and described

(Quote 5) the (fire) - 'fire' has been mentioned previously, so we know we are talking about the same fire

(Quote 6) the (hotel) – refers to the same hotel as mentioned before

(Quote 7) the (hotel) – identifies the hotel and refers back to the Grand Hotel

'Specificity' was also identified when given references were associated with NPs earlier introduced in the discourse. In the examples below, the NS students used *the* in front of the associative anaphoric NPs *hotel elevators, hotel manager*, and *situation*, as these NPs exploited the shared knowledge of their relations with the previously-mentioned nouns *hotel* and *fire*.

(Quote 8) the (hotel elevators) elevators belong to or are part of the specific hotel

(Quote 9) the (hotel manager) – manager of the same hotel we have been reading about

(Quote 10) the (situation) – identifies the 'situation' and refers to the event of the fire described in the text

(Quote 11) the (situation) – makes reference to the specific fire incident

In the above contexts, identification of the referred NPs was possible for the hearer in the endophoric uses (i.e. from the linguistic cues provided in the text), and thus the definite article *the* was required in the given contexts.

From the NS metalinguistic explanations, it was found that there were some native speakers who explained their choice of *the* based on the plural form of given NPs such as *guests* and *cats*. Theoretically, for definite descriptions the distinctions of countability and number are neutralized. In other words, *the* marks a count use of a noun in both a singular or plural form. Thus, plurality is not related to the definite concept. The key feature that calls for *the* is when a given reference fits in the mutual knowledge between the speaker and the hearer. At this stage, it may be too early to make an assumption that the native speakers formulated a theory that 'a plural noun calls for *the*'. Indeed, it is not clear at this point as to whether the native speakers in this study also considered 'referentiality' as part of their justification for using *the* but did not mention it explicitly, or they related *the* solely with a plural form of a noun. This particular issue may require further research.

Another notable point in the NS students' justifications for using *the* was that the term 'specific' was used extensively in a large number of explanations. Although the notion of HK was referred to by some students in comments such as "*the fire has already been talked about previously,"* or *"the hotel has already been mentioned*," it appeared that many of them did not understand the relationships that exist between 'specific reference' (SR) and 'hearer's knowledge' (HK). The two notions were often found to be compressed into one general concept. Thus, the reference that the native speakers considered to be 'specific' was usually treated as if it was identifiable by the hearer [+HK]. Accordingly, the term 'specific' was used to a great extent when the native speakers chose to use *the*.

There were cases, however, when the term 'specific' was used to represent [+SR] by some native speakers. In such cases, a referred NP was considered to be identifiable to the speaker, not the hearer, as in the items in Type 3 [+SR, -HK] NPs marked with the indefinite article *a/an*. Therefore, we also see the term 'specific' alternatively used in comments that justified the choice of *a/an* in Type 3 contexts, as in the examples below:

(Quote 12) a (police car) - singular, specific car

(Quote 13) a (man) - specific, singular man

Other NS students expressed their lack of understanding that the notion of SR should be separate from HK. These native speakers, thus, used the term 'specific' to represent [+SR] at one time and [+HK] at the other. This means that the same term—i.e. 'specific'—was alternatively used to explain their choice of *a/an* in some items in Type 3 and also their choice of *the* in some items in Type 2. For example, one student explained her choice of *a* in a first-mention NP *fire* as: "a - one singular, *specific fire*" and in her choice of *the* when *fire* appeared again as: "the - referring to *the specific fire*."

(2) Use of *a/an* in definite (Type 2) contexts

As shown in Table 6.4 above, the indefinite article *a/an* was used instead of *the* in only one token (0.3%). Theoretically speaking, when native speakers do not use the expected article, one cannot claim that their article choice is wrong. Based on the metalinguistic explanations provided in the study, the native speakers' variable choices were found to result from their different intuitive evaluations or conceptual impressions of the contexts. In the case of using *a* instead of *the* in front of *thrill*, for example, one native speaker considered the given noun to mean "only one thrill" being referred to instead of *the thrill* that was identifiable by the hearer.

(3) Use of Ø in definite (Type 2) contexts

The use of Ø instead of *the* in Type 2 NPs was found in 12 tokens (3.3%) in the NS group. Of the 12 tokens of Ø use, 8 tokens were supplied with non-explicit explanations. Pragmatic (P) explanations (i.e. the NPs were not 'specific') were given to 3 uses of Ø and the remaining 1 use of Ø was explained based on a semantico-syntactic (S) reason (i.e. the NP was plural).

It should be noted that of the twelve uses of Ø for the target article *the*, eight tokens were found in item 38: '…one of several of (38) *the* large cats *that are graceful but powerful killers*…' (Please refer to Table 6.1). Among these eight variable uses of Ø in this item, five native speakers did not explain explicitly why they thought Ø was appropriate, while the remaining three native speakers indicated that Ø was used because

the referred NP did not need to be specified. As can be seen, the existence of a post-modifying relative clause: '*that are graceful but powerful killers*' was intended to serve as an indicator for the reference of *large cats* to be identifiable by the reader. However, there is some complexity in the structure of the referred NP. The structures 'several + Ø NPs' and 'several of + the NPs' are both grammatical, as in:

'…one of several Ø large cats that…' or
'…one of several *of the* large cats that…'

Therefore, it might be possible that some native speakers failed to recognize the preposition *of* after *several*, and thus, they chose to use Ø instead of *the*. This item may be an inappropriate representation of Type 2 context and an adjustment to this test item may be necessary.

6.2.1.2 Metalinguistic Explanations of the NNS-H Learners in the Use of *the* in Definite Contexts

As demonstrated in Table 6.3, the NNS-H students performed best with the definite article *the* in Type 2 (referential definite) contexts. The rate of target article use in Type 2 NPs reached 73.1%. From the quantitative analysis of MLK presented in Table 5.8 on page 102 of Chapter 5, the NNS-H students explained their article choices for Type 2 NPs based primarily on pragmatic MLK in the same way the NS students did. The following summary reports on the NNS-H students' metalinguistic explanations for their use of the target article *the*, followed by their uses of *a/an* and Ø.

(1) Use of the target article *the*

The metalinguistic explanations given by the NNS-H learners in justifying their use of *the* were consistent with those offered by the NS students. Four main factors were used to explain their choice of the target article *the*:

1.1 a previously-mentioned NP
1.2 the existence of a modifier (e.g., an adjective or superlative)
1.3 the association of a given referent with a previously-mentioned NP
1.4 the plural form of a given noun

The rationale that ‘a given NP has already been mentioned’ was used most frequently, particularly with the previously-introduced nouns: *hotel* and *fire*. The test results showed that all thirty students correctly chose *the* for *hotel*; however, for the noun *fire*, *the* was chosen by only twenty-one students. As far as can be seen from the students’ explanations, one factor that accounted for their error had to do with the detection of noun countability. As one can see, *hotel* tends to be more easily treated as countable, while *fire* is more difficult to judge, for it can be used as countable and uncountable in different contexts. The students who considered *fire* to be uncountable in this context chose to use *Ø* instead of *the*. For Thai EFL learners, noun countability appears to be a prominent factor in their article use and it poses numerous problems in making proper article choices for them.

The second reason for the NNS-H students in using *the* was when there was a modifier attached to a given noun. With the existence of modifiers, the students often explained their article choices based on (S) MLK using local structural cues rather than considering global contextual information. This can be seen in comments for the NPs with the adjectives *best*, *same*, and *first*.

(Quote 14) the (best thinker) – the superlative adjective calls for ‘the’

(Quote 15) the (same way) - ‘same’ needs ‘the’ to identify which ‘way’

(Quote 16) the - ‘first’ tells the order and must be preceded by ‘the’.

When a reference is associated with a previously-mentioned object or event, it is assumed to be known by the hearer. Like the NS group, the NNS-H students explained their choice of *the* in the NPs *hotel elevators, hotel manager*, and *situation* based on this rationale. Explanations included statements such as:

(Quote 17) the (hotel elevators) - we already know which *hotel*

(Quote 18) the (hotel manager) - *manager* is specific because it refers to the manager of the specific hotel

(Quote 19) the (situation) - which *situation* has been identified

Another commonly-observed theory that many NNS-H students used to explain their choice of *the* was a non-generalizable rule that “*the* should be used with a plural noun.” This rationale was also used by some native speakers as discussed in the

previous section. A number of NNS-H students used this explanation to justify their use of *the* in front of the plural NPs *guests* and *large cats*. The non-generalizable rule that 'plural nouns take *the*' may have been formulated on a strong notion that '*a/an* is used with a singular noun'; thus, *the* should be used with a plural noun.' It was found that this hypothesis was commonly observed among a large number of NNS students in both proficiency groups.

It was notable that in justifying the use of *the*, the term 'specific' was also used to represent SR and HK in the same way as it was used by the NS participants. However, it was found that the observed number of the NNS-H students who explicitly referred to HK in their choice of *the* was higher than the number of the native speakers who mentioned HK in their explanations. It appeared that the NNS-H students began to realize that HK was one of the key features of referentiality in determining articles. This is probably due to the fact that article usage is often taught through a few basic rules that EFL learners are supposed to follow when they have to make article choices. Examples of such rules include:

(Rule 1) When a noun is countable and singular, 'a' is used.

(Rule 2) 'a' is used only when the following word begins with a consonant (sound). If the following word begins with a vowel (sound), then 'an' is used.

(Rule 3) When a noun is introduced for the first time, 'a' should be used, but when the same object or event is mentioned for the second time, then 'the' should be used.

(Rule 4) When an object or event is not specific, 'a' is used. If it is specific, 'the' is used.

Concepts such as '*first-mention'* and '*second-* or *subsequent-mention*' are often referred to when the teacher explains the components of the English article system in ESL/EFL classes. As metalinguistic (explicit) knowledge is normally gained through explicit learning from teachers or textbooks, the reasons articulated by the learner are likely to reflect the type and amount of input (i.e. language instruction) the learner has received. Given the fact that the input is readily abundant in a natural setting of L1 for native speakers, it may be hypothesized that the amount of formal instruction in article usage is likely to be far less than the amount of article instruction offered in EFL lessons.

This explains why statements such as ‘we know…’ ‘it is known…’ and ‘previously mentioned’ were fewer in the explanations offered by the NS students than in the reasons given by the NNS-H students in the study.

Although the students had often learned how to use articles through some basic rules, they still had problems with article usage. It was found that application of these seemingly simple rules was not as simple as they seemed to be. A result of the students’ efforts in using articles was that they had to form various hypotheses or non-generalizable rules, including collocations, regarding article use. Most hypotheses were syntactically-based and were inconsistent with linguistic theory. The following are some examples of non-generalizable rules found in the NNS-H students’ explanations (more examples are shown in Table 6.8, which appears at the end of the chapter).

1. Non-generalizable collocational rules:
   1.1 Certain prepositions with certain articles
      e.g. “at + the,” “on + *Ø*”
   1.2 Certain articles with certain nouns
      e.g. “the + hotel,” “*Ø* + manager”

2. Non-generalizable hypotheses:
   2.1 An adjective + certain articles
      e.g. “a + adjective,” “the + adjective”
   2.2 No article follows a preposition
      e.g. “preposition + *Ø*”

Most of the students’ non-generalizable rules were found with their choices of *the* and *Ø* rather than with the choice of *a/an*. Generally, when the students used *a* or *an*, they usually stated a few basic rules they had presumably learned in EFL classes, such as “the noun is singular,” “the reference is not specific,” or “the noun is mentioned for the first time.”

(2) Use of *a/an* in definite (Type 2) contexts

The use of *a/an* instead of *the* in the NNS-H group was explained based on the following factors:

2.1 the singular form of a given noun

2.2 a 'non-specific' NP

2.3 a first-mention NP

One crucial reason the NNS-H students chose *a/an* instead of *the* was due to the failure to detect the lexical cohesion that existed between words (such as synonyms). For example, many students did not understand that the word *prey* was used to refer back to the word *animal (being hunted)* mentioned previously. Thus, they perceived *prey* as a first-mention noun and *a* was used instead of *the*. This suggests that students need to have sufficient lexical knowledge in order to consider contexts dynamically.

(3) Use of Ø in definite (Type 2) contexts

Basically, the NNS-H students justified their use of Ø (instead of *the*) with the reason that the NP was plural or that the reference in question was considered to be 'non-specific'. Another justification for the use of Ø was also found to be caused by the students' misreading of a word or misinterpretation of the meaning of a word. This includes a lexical misjudgment of a word class. Examples of this are illustrated in the explanations for the choice of Ø in front of *thrill*, used as a noun in its given context.

(Quote 20) Ø (thrill) – 'thrill' is a verb; no article is needed

(Quote 21) Ø (thrill) – 'thrill' is an adjective

Other uses of Ø were explained based on the students' non-generalizable hypotheses including collocations of 'Ø + a certain noun' or 'Ø + a certain preposition'. It was also found that the students supplied more non-explicit reasons with their use of Ø than with *a/an* or *the*.

To sum up, the MLK used in explaining the choice of *the* in definite contexts in the NNS-H group was similar to the MLK used by the NS students. Both NS and NNS-H groups used MLK that was more pragmatically-oriented and they often associated the use of *the* with the term 'specific'. They judged the given references to be 'specific' based on similar factors. However, the NNS-H students often alluded to the instructional rules and used words such as 'second or previously-mentioned' or 'known' to justify their use of *the* more often than the native speakers. This may be because the basic

article rules provided in most EFL classes often use these words to describe the notion of HK. Despite referring to the HK notion in their explanations, the NS and NNS-H students did not seem to have a clear understanding of the distinction between HK and SR and often compressed these two notions into one general concept under the commonly-used term 'specific'.

6.2.1.3 Metalinguistic Explanations of the NNS-L Learners in the Use of *the* in Definite Contexts

As a whole, the NNS-L students used *a/an* more than *the* for all NP types. In Type 2 contexts where the definite article *the* was the target article, *the* was used in only about half of the responses (53.9%) while *a/an* was chosen instead of *the* in about one-third of the responses (34.2%) (Please refer to Table 6.4 above). I will summarize the metalinguistic explanations given by the NNS-L students in their use of *the, a/an* and *Ø*.

(1) Use of the target article *the* in definite (Type 2) contexts

In using articles, it was found that the NNS-L students relied heavily on local structural cues (S-MLK) rather than referentiality (P-MLK). Even in the use of *the*, the referentiality features SR and HK were seldom referred to in their explanations. This could be due to the fact that considering referentiality requires the students' ability to comprehend contexts. As the NNS-L students had limited lexical knowledge, they lacked the ability to consider larger contextual dynamics. Their article judgments, thus, had to rely on limited structural cues instead.

Of the two components of referentiality, the notion of HK was found in the students' explanations in a smaller number than the notion of SR. This gives evidence that for the NNS-L students, detection of HK was more problematic than SR. Most NNS-L students still failed to regard HK as the primary requirement for choosing *the* as opposed to *a/an*. They did not realize that specific reference alone does not denote the use of *the* if the reference is not assumed shared knowledge between the speaker and the hearer.

There was evidence that, in making article choices, most of the NNS-L students were mainly constrained by a few basic rules often provided in most EFL article lessons.

From the examples of instructional rules given in the previous section, the rules prescribed for the use of *the* (i.e. Rule 3 and Rule 4) appear to be more context-oriented. The students accordingly had difficulty applying these rules to identify the conditions for the proper use of *the*. As a result, to accommodate the use of the definite article, the NNS-L students had to form various non-generalizable hypotheses, including collocations. The students' invented rules differed probably due to varying amounts and types of input that the students had been provided. Some rules were commonly observed across students, while others were more individually-based. Examples below illustrate some of the students' non-generalizable hypotheses used to explain the choice of *the* in Type 2 context.

1. Rules that were commonly-observed among a number of students:
   1.1 '*the* + plural nouns (or nouns with –s)'
   1.2 '*the* + adjective'
   1.3 '*the* + nouns referring to places'

2. Rules that were individually-based:

   2.1 'on + *the*'
   2.2 '*the* + nouns referring to objects'
   2.3 '*the* + non-count nouns'
   2.4 '*the* should begin a sentence'
   2.5 '*the* + nouns in the subject position'
   2.6 '*the* is used at the end of the sentence'

The students' explanations suggested that rules that were commonly-observed among the students tended to be applied more consistently than rules that were individually-based. Individually-based rules were generated depending on the student's different types and amounts of input and varying degrees of attention to the input. The use of these rules was found to differ from context to context. Moreover, many NNS-L students expressed their confusion over the application of these rules. Some students applied the same rule to different contexts, while others applied two different rules to the same NP environment.

(2) Use of *a/an* in definite (Type 2) contexts

As one can see, the NNS-L students used *a/an* instead of *the* in Type 2 context in 34.2%. This means that they overused *a/an* substantially in the definite context. Why was *a/an* overused in the low-proficiency group?

As can be seen from the examples of instructional rules described in the previous section, the use of *the* (as stated in Rule 3 and Rule 4) is context-dependent; it requires the student's ability to comprehend larger dynamic contexts, which is often lacking in students at a low-proficiency level. The proper use of *the*, thus, appears to be problematic for low-proficiency students. In contrast, the few rules for determining the indefinite article *a/an* (i.e. Rule 1 and Rule 2) are more syntactically-based. Also, components such as singularity seem to be more easily observable and less complicated than referentiality features. As a result, in a task that requires article use, students with low proficiency are likely to initially consider singularity or countability of a given noun and decide to use *a* (or *an*) first if the conditions for Rule 1 and Rule 2 pertain.

Another reason why *a/an* was overused among the NNS-L students might be that the general notion of the English indefinite article seemed to be less complex than the use of *the* in English definite NPs. The typical use of *nɣ̀ŋ* ('one') in Thai indefinite descriptions (Refer to Section 2.2.1.2 in Chapter 2) is somewhat equivalent to the use of *a/an* in English. It may be true that the rigidness of using *nɣ̀ŋ* and *a/an* is different, for *nɣ̀ŋ* is not always obligatory in Thai in the same situations in which it is in English. Despite this fact, drawing an analogy between *nɣ̀ŋ* and *a/an* in indefinite contexts seems to be less difficult than mapping the linguistic realizations of definite descriptions between the two languages. As Thai is a discourse-oriented language, a definite marker is often considered redundant (Refer to Section 2.2.1.1 in Chapter 2), whereas in English, *the* is obligatory for definite descriptions.

As article usage is often taught through a few basic rules, most NNS-L students were often constrained by rules and based their judgments on a rule-based hypothesis, either consistent or inconsistent with linguistic theory. The students usually exercised rules independent of each other. The following are examples of rules cited by the NNS-L students for their choice of *a/an*.

1. Rules that were consistent with linguistic theory:

    1.1 '*a* + countable nouns'
    1.2 '*a* + singular nouns (or nouns with no –*s*)'
    1.3 '*a* + nouns beginning with a consonant'
    1.4 '*an* + nouns beginning with a vowel'
    1.5 '*a* means one'

2. Hypotheses that were not consistent with linguistic theory:

    2.1 '*a* + nouns referring to a person'
    2.2 '*a* + non-count nouns'
    2.3 '*a* + common nouns'
    2.4 '*a* + adjective'
    2.5 'of + *a*'
    2.6 '*a* + nouns in the object position'
    2.7 '*an* is used for connecting words'

(3) Use of Ø in definite (Type 2) contexts

In the NNS-L group, Ø was used instead of *the* due to the students' failure to identify a head noun or misjudgment of a word class. For example, two students wrote:

(Quote 22) Ø (thrill) – no article because 'thrill' is a verb
(Quote 23) Ø (guests) – because 'guests' is not a noun

In the students' use of Ø, it was found that they had more difficulty articulating a reason for using this article than for their choices of *the* and *a/an*. While the NNS-L students attempted to formulate rules for using *the* and *a/an* and explained their choices explicitly, they often expressed uncertainty when they chose to use Ø. Thus, a number of their explanations were unclear; for example,

(Quote 24) Ø – nothing seems to fit here
(Quote 25) Ø – I guess Ø might be correct
(Quote 26) Ø – I don't know which article to use
(Quote 27) Ø – it sounds better without an article

(Quote 28) Ø – the sentence is already complete

As indicated earlier, the use of Ø constituted the lowest proportion of article uses among the NNS-L students. Moreover, the students had more problems providing an explicit reason to support their choice of this article. The difficulty in properly using Ø and articulating a reason for why they chose this article might be due to the fact that Ø compiles too many complex semantic, syntactic, and pragmatic functions, which include referentiality features (SR, HK), countability and number. This seems to explain why manipulating these multiple functions for the proper use of this article leads to a considerable burden for the students, particularly those at a low-proficiency level.

So far, we have seen that the NNS-L students' MLK in using and explaining their article choices in definite contexts was different from the MLK of the NS group. The NS students had a high rate of targetlike use of the definite article and they explained their choices primarily with pragmatic MLK. The NNS-H students' use of MLK was more consistent with the native speakers than the NNS-L students, particularly in using pragmatic MLK to explain their target article use. Due to limited lexical knowledge, the NNS-L students were more constrained by rules and relied heavily on local structural cues in making article choices. The students used *a/an* substantially based on semantico-syntactic MLK by referring mainly to countability and the singular form of a given noun. Pragmatic MLK was seldom used by the NNS-L students.

In the following sections, the metalinguistic explanations given by the participants in their article choices in indefinite contexts will be summarized and discussed. Indefinite noun phrases are marked by two articles: *a/an* and Ø. In Huebner's NP classifications, these two articles mark Type 3 [+SR, -HK] referential indefinites and Type 4 [-SR, -HK] non-referentials. One common component shared by both NP types is that a referred NP is unidentified by the hearer—i.e. the reference is not yet a part of the assumed hearer's knowledge [-HK]. In Type 3, the NP has a specific reference [+SR], and it is often called a first-mention noun. The referred NP in Type 4, in contrast, does not have a specific reference [-SR], and it is referred to as non-referential use. I will discuss each of the two articles in turn. First, I will begin with the overt form of the indefinite articles: namely, *a/an*. Then, the *zero* article (Ø) will be discussed.

6.2.2 Metalinguistic Explanations in the Use of *a/an* in Indefinite Contexts

As I mentioned in Chapter 3, the indefinite article *a/an* marks [[+count][+sing]] NPs in Type 3 and Type 4. Table 6.5 below presents the distributions of article choices (with percentages in parentheses) performed by the three sample groups in Type 3 and Type 4 contexts marked with the indefinite article *a/an*.

Table 6.5: *Frequency Distribution and Percentage of Article Uses in Type 3 and Type 4 Marked with a/an*

Type		Type 3 [+SR -HK]			Type 4 [-SR -HK]		
Target Article		a/an (180 tokens)			a/an (180 tokens)		
Article Choice		Ø	a/an	the	Ø	a/an	the
Group	NS	1 (0.5%)	**172 (95.6%)**	7 (3.9%)	1 (0.5%)	**156 (86.7%)**	23 (12.8%)
	NNS-H	17 (9.4%)	**124 (68.9%)**	39 (21.7%)	23 (12.8%)	**129 (71.7%)**	28 (15.6%)
	NNS-L	10 (5.6%)	**110 (61.1%)**	60 (33.3%)	26 (14.4%)	**108 (60.0%)**	46 (25.6%)

In Type 3, the average score in using the target article *a/an* in the NS group was substantially higher than the scores of the two NNS groups. The differences of scores between the NS and the NNS groups are statistically significant. The gap was smaller in Type 4 context, but the differences in performance between the NS and NNS groups are still significant. Of the two NNS groups, if one recalls, the scores in both Type 3 and Type 4 marked with *a/an* did not reveal a statistically-significant difference (please refer to Table 5.6 on page 97 of Chapter 5). In the following sections, I will first present metalinguistic explanations of the NS group with regard to the use of *a/an* in indefinite contexts. Then, explanations of the Thai learners are compared across proficiency groups so that the problems underlying the use of *a/an* in each group can be identified.

6.2.2.1 Metalinguistic Explanations of the Native Speakers in the Use of *a/an* in Indefinite Contexts

(1) Use of the target article *a/an* in indefinite (Type 3 and Type 4) contexts

In using *a/an* in indefinite contexts, the NS explained their choice based on four main factors:

1.1 a ‘non-specific’ NP

1.2 the singular form of a given noun

1.3 the phono-syntactic rule of the distinction between the consonant/vowel sound of the following word

1.4 a ‘generic’ interpretation of a given NP (mainly for items in Type 4)

For the choice of *a/an* in Type 3 first-mention use, referentiality was commonly referred to in the metalinguistic explanations of the NS group. As noted in the previous section, when referring to a referentiality feature, the NS students often alternatively used the term ‘specific/non-specific’ to refer to the notion of HK, SR, or both. The use of one term to refer to two notions was evident when the NS students explained their choices of *a/an* in Type 3 items as illustrated in the following examples:

(Quote 29) a (fire) – no specific fire stated yet

(Quote 30) a (fire) – you don’t know specifically what fire yet

(Quote 31) a (fire) – ‘fire’ is not specific

(Quote 32) a (fire) – describes a specific singular ‘fire’

(Quote 33) a (fire) – it is talking about a specific fire

(Quote 34) a (fire) – one singular, specific fire

In comments 29 - 31, the native speakers used ‘specific’ to denote HK. Thus, when they commented that *fire* was ‘not specific,’ they probably regarded this ‘first-mention’ NP to be non-identifiable by the hearer [-HK]. On the contrary, the NS students who remarked that *fire* was ‘specific’ (as in comments 32 - 34) may have used this term to indicate that the reference was identifiable to the speaker [+SR].

There were also some students who used ‘specific’ to refer to [+SR] at one time and [+HK] at the other. For example, one student explained his choice of *a* in the first-mention noun *fire* as follow:

(Quote 35) a (fire) – it is a specific event

When *fire* was mentioned for the second time in the text, he explained his choice of *the* in this instance as:

(Quote 36) the (fire) – the specific event

In using *a/an*, a linguistic rule such as 'a given NP is singular' was also mentioned by a number of NS students. With the use of *an*, the phono-syntactic rule of the following word beginning with a consonant/vowel sound was cited in a large number of explanations. Some NS students referred to this rule independently, while others incorporated this rule with explanations on referentiality.

In Type 4 non-referential use, the indefinite article *a/an* marks a NP whose referent is non-identifiable by both the speaker and hearer [–SR –HK]. Thus, the reasons given by the NS students to explain their uses of *a/an* included words such as "non-specific," "general," "one of many," "any," or "a representative of the whole." For example,

(Quote 37) a (vessel) – the subject can be *any* vessel in that time
(Quote 38) a (vessel) – speaking of a *generic* vessel dolphins swim alongside
(Quote 39) a (rock) – this sentence is talking about a rock *in general*
(Quote 40) a (rock) – a rock is *a representative of the whole*
(Quote 41) a (lion cub) – shows *generalization for 'all'* lion cubs born
(Quote 42) a (lion cub) – using singular nouns to talk about *all* lion cubs

As noted earlier, the non-referential use and the generic use are difficult to distinguish one from the other. It is, therefore, not uncommon to find that many native speakers who considered using *a/an* in items in Type 4 based their choice on a generic reading of the context.

(2) Use of *the* in indefinite (Type 3 and Type 4) contexts

The use of *the* in indefinite contexts among the NS students was found to be based on two main reasons:

2.1 a 'specific' NP
2.2 a 'generic' interpretation of a given NP

The choice of *the* in Type 3 (first-mention) context was mainly due to the NS students' consideration that a given NP was 'specific.' For example, *party* in '…a number of guests who were enjoying (26) *a* party at the hotel…' was considered to be 'specific' as explained in the following comments offered by those who used *the*:

(Quote 43) the (party) – identifies the 'party' that was taking place at that time

(Quote 44) the (party) – describes which 'party'

(Quote 45) the (party) – talking about the specific 'party' that is going on

Indeed the word *party* was introduced for the first time. Why, then, did some native speakers consider the NP in question to be 'specific'? It might be possible that the NS students regarded the post-modifying phrase "at the hotel" as a cue that the referred NP should be 'specific." If one considers 'party' to be the only party at this particular 'hotel', then 'party' should be judged as identifiable by the hearer. This, of course, is not necessarily true.

Another example is the use of *the* in front of *man* in '…(32) A man told our reporter that…'. In this case, there was no modifier attached to the word *man*, which was introduced for the first time in the passage. Yet, a few native speakers chose to use *the* for this given context. The insertion of *the* in front of *man* came from a similar reason:

(Quote 46) the (man) – describes *which man* and helps the sentence flow

(Quote 47) the (man) – refers to a *specific man*

From the above explanations, it was not clear as to why *man* was considered to be 'specific' by those who used *the* as it was a first-mention use and it is difficult to see the connection of this word to any other words in the text.

A similar reason was also given by another student for his choice of *the* in the NP *police car*:

(Quote 48) the (police car) – a specific car

Again, in this context *car* was mentioned for the first time in the text. The difference is that *car* is modified by the word *police*, while *man* has no modifiers. However, the student did not mention in his comment that it was the word *police* that made the reference of *car* 'specific'. As *the* was used by only a few native English speakers for each of the above test items, it limits the ability to draw inferences from the evidence at this stage.

As indicated in Table 6.5, *the* was chosen instead of *a/an* at a greater rate in Type 4 (12.8%) than in Type 3 (3.9%) by the NS group. The use of *the* to replace the target article *a/an* in Type 4 (non-referential) contexts was due to two factors. The first factor was consistent with the use of *the* instead of *a/an* in Type 3 items as discussed above. References in some items in Type 4 were also considered to be 'specific' as was the case of the use of *the* in some Type 3 NPs. For example, eight NS students who used *the* in front of *vessel* explained their choice based on this reason.

(Quote 49) the (vessel) – refers to a specific vessel

Another reason given by the native speakers to explain the choice of *the* in some Type 4 items was due to the interpretation of a given NP in a 'generic' sense. For example, related to the same item above, one student explained his choice of *the* for the noun *vessel* by saying that it referred generically to 'any type of vessel'. Likewise, another student chose *the* in front of the noun *animal* and explained that:

(Quote 50) the (animal) – 'the' is used because it can be 'any' animal as her prey

As one might expect, the distinction between generic descriptions and non-referential NPs can be very difficult. A non-referential NP and a generic use can both be interpreted as 'any' entity that represents many or all entities sharing the same characteristics. This is supported by the explanations given by the NS students in this study that some NPs in their non-referential use may allow a generic reading. When a NP is conceived as generic, the article *the* is possible to formally convey generality of a given reference as a whole set.

Since the replacement of *the* in Type 4 *a/an* context was based on more than one ground (i.e. the consideration of a given reference as assumed known to the hearer and as a generic interpretation), this might explain why *the* was used instead of *a/an* in a higher number in Type 4 than in Type 3.

(3) Use of Ø in indefinite (Type 3 and Type 4) contexts

As demonstrated in Table 6.5, the NS group chose Ø instead of *a/an* in only one token each in Type 3 and Type 4. Theoretically, the article *a/an* contrasts with Ø in the aspects of countability and number. A singular, count noun calls for *a/an*; whereas a

plural, count noun or a mass noun takes Ø. Thus, in a context where *a/an* is required, the choice of Ø should appear to be inappropriate. Each token of Ø in each NP type was chosen by the same student who did not offer an explicit reason as to why he chose Ø for each environment.

In the following two sections, I will report on the metalinguistic explanations of the NNS-H and NNS-L students in their uses of articles in the indefinite *a/an* contexts. As demonstrated in Table 6.5, the NNS-H group did not perform better significantly than the NNS-L group in the use of *a/an* in Type 3 and Type 4. Statistically, the significant difference in performance between the two groups was not found at the alpha .05 level, as reported in Table 5.6 on page 97 of the previous chapter. In order to understand the problems underlying the use of *a/an* in each group, I first explore metalinguistic explanations for article choices among the NNS-H students, and then I analyze explanations that supported article uses in the NNS-L group in the section that follows.

6.2.2.2 Metalinguistic Explanations of the NNS-H Learners in the Use of *a/an* in Indefinite Contexts

(1) Use of the target article *a/an* in indefinite (Type 3 and Type 4) contexts

Roughly speaking, the reasons given by the NNS-H students for their use of the target article *a/an* were somewhat consistent with linguistic theory although a majority of the explanations contained only a partial description of the features required for the proper use of *a/an* in a given context. Generally, the NNS-H students referred to three main factors that supported their choice of *a/an*:

1.1 a first-mention or 'non-specific' NP (i.e. referentiality)
1.2 a countable noun
1.3 the singular form of a given noun

In Type 3 (first-mention) context, it was found that the students concentrated more on referentiality than the other features. This might be due to the fact that nouns and noun phrases representing Type 3 are often used in narrative texts. In order for the students to capture the flow of a narrated event, SR and HK features appeared to take

priority over the other aspects. Common explanations used among the students in the NNS-H group to support their choice of *a/an*, thus, included statements such as:

(Quote 51) a (fire) – this noun is firstly-introduced

(Quote 52) a (party) – we don't know which party is being referred to

(Quote 53) a (German woman) – 'German woman' has not been talked about before

(Quote 54) a (police car) – which police car is not specified

The above explanations indicate that the NNS-H students were able to figure out that the feature [-HK] was the primary requirement that determined the article *a/an* as opposed to *the*.

(2) Use of *the* in indefinite (Type 3 and Type 4) contexts

The choice of *the* in the NNS-H group constituted 21.7% and 15.6% of article choices in Type 3 and Type 4 respectively. The NNS-H students chose *the* instead of *a/an* in indefinite contexts based mainly on referentiality. The students' explanations indicated that they considered given references to be 'specific'. According to the NNS-H students, references were judged to be 'specific' based on:

2.1 the existence of modifiers

2.2 their own assumptions about the context

Among the NNS-H students, the existence of modifiers was commonly regarded as a device to make references 'specific' and identifiable by the hearer. Thus, with a modifier attached to the referent in NPs such as *empty gasoline can, German woman*, and *police car*, these NPs were judged to be definite. The following are some examples:

(Quote 55) the (gasoline can) – refers to the can containing gasoline, not any other cans

(Quote 56) the (German woman) – it is specific, considering the word 'German' which modifies 'woman'

(Quote 57) the (police car) – specifically referring to 'car' of the police

When the students treated modifiers such as adjectives as a device to indicate definiteness (as in the above examples), these students often wrongly used *the* in contexts where *the* was not supposed to be used. As a matter of fact, having a modifier does not necessarily mean that the reference is identifiable by the speaker; needless to say, it might not be identifiable by the hearer.

Another factor responsible for the students' inaccurate uses of *the* is that some students brought too many assumptions to their readings. It was found that when a referred noun was related in some way to a word or phrase that appears in the text, the students often wrongly judged the reference as if it was known to the reader. Thus, *the* was thought to be required. An example that illustrates this is when some students related the first-mention noun *fire* with the time expression *yesterday evening*, and explained their uses of *the* in front of *fire* as below:

(Quote 58) the (fire) – it specifically refers to the 'fire' that broke out *yesterday evening*

In some cases, the students used a word or phrase that appears later in the text as a device that marks definiteness. Related to the same item mentioned above, a few students who used *the* explained that *fire* (in '…(20)___fire badly damaged the Grand Hotel in San Bernardino…') was identifiable by the hearer.

(Quote 59) the (fire) – we should know that it is talking about the 'fire' at the Grand Hotel.

Indeed, in a common reading, when *fire* was firstly-introduced, it was not yet a part of the assumed hearer's knowledge. Thus, *the* is not meant to be used here.

Some students regarded the word *police* in *police car* and *lion* in *lion cub* as an indicator that the head noun *car* and *cub* called for *the*. In their comments, they explained that *police* and *lion* had already been mentioned; thus, with the presence of these words as a modifier for *car* and *cub*, the references in question should be assumed known to the hearer.

(Quote 60) the (police car) – because 'police' has been talked about before

(Quote 61) the (lion cub) – because 'lion' was mentioned in the previous paragraph

One may recall from the quantitative analysis reported in Chapter 5 that the MLK of the NNS-H students were primarily pragmatically-oriented. These students began to realize that, in order to use *the* as opposed to *a/an*, the reference should be identifiable not only by the speaker but also by the hearer. Thus, they had to consider contexts more seriously and dynamically. However, correctly detecting whether or not a reference should be identifiable by the hearer turned out to be difficult for the Thai students. As they became more proficient, the NNS-H students more often wrongly judged references to be part of the assumed hearer's knowledge than was actually the case. By incorporating too many presuppositions to their readings, such assumptions appeared to interfere with their accurate detection of the assumed hearer's knowledge. References were often taken for granted to be identifiable by the hearer, and thus *the* was overused by the high-proficiency students. The overuse of *the* in Type 3 and Type 4 contexts decreased the accuracy rate for the target article *a/an*. This may be a possible reason why the performances in these two NP contexts between the students in the two proficiency groups did not yield a statistically significant difference.

(3) Use of Ø in indefinite (Type 3 and Type 4) contexts

As nouns and noun phrases that require the article *a/an* are countable and singular, the *zero* article (Ø) is inappropriate to mark these NPs. Among the NNS-H students, Ø was inaccurately used in Type 3 and Type 4 in 9.4% and 12.8% respectively, as displayed in Table 6.5 on page 133. One common reason found among the NNS-H students who chose Ø instead of *a/an* had to do with the misjudgment of noun countability. Some students, for example, explained their choice of Ø in front of *fire* and *rush* based on their consideration that both nouns were uncountable.

It was also found that a large number of Ø uses were based on non-generalizable hypotheses and distorted collocational rules as well as impressionistic judgments and guessing. Such explanations indicated that Thai students, even with high English proficiency, did not have a clear understanding of *zero* article usage. The students' difficulty in using Ø will be discussed in more depth in Section 6.2.3.2.

To sum up, the use of MLK in determining the article *a/an* in indefinite contexts between the NS and NNS-H groups was similar in that both groups relied on pragmatic (P) MLK as well as (S) MLK. Referring to Table 5.9 on page 106 of the previous chapter, one can conclude that the NS students explained their choice of *a/an* based on the singular form of given NPs, referentiality, or the combination. The NNS-H group also chose *a/an* based on the singular form of NPs, countability, referentiality, or the combination (PS) MLK. In Type 3, the NNS-H students used (P) MLK more often in the narrative text. In their efforts to consider the context dynamically, they brought too much assumption to their readings, and thus references were wrongly judged to be part of the assumed hearer's knowledge. As a result, *the* was used more than it was intended to be used in the indefinite context. The explanations provided by the NNS-H students supported the results in Table 5.9 that the NNS-H group used (P) MLK in a higher percentage than the NS group in Type 3 marked with *a/an*.

6.2.2.3 Metalinguistic Explanations of the NNS-L Learners in the Use of *a/an* in Indefinite Contexts

We have seen in the above section why the high-proficiency students overused *the* in *a/an* contexts. In what follows, we will explore the metalinguistic reasons responsible for the target article uses as well as errors in the NNS-L group in indefinite *a/an* contexts.

(1) Use of the target article *a/an* in indefinite (Type 3 and Type 4) contexts

For the choice of *a/an* in the NNS-L group, the explanations were more consistent with linguistic theory than with their choices of the other articles. The NNS-L students mainly referred to two factors, considered to be (S) MLK:

(1) the singular form of a given noun

(2) a countable noun

As earlier mentioned, article usage is often introduced in EFL classes through basic rules (Please refer to Section 6.2.1.2). Each rule usually provides a description of one feature underlying the use of each article. Normally, in order to determine a proper article, more than one rule is necessary. This means that the students need to have a

clear understanding of the relationships that exist among certain rules when making article choices. For example, for a student to use *a/an* appropriately, Rules 1, 2 and 3 have to be applied simultaneously. As far as their explanations indicated, the NNS-L students had problems with exercising these instructional rules. They often exercised each rule independently without having a clear understanding of their interdependence with each other. This was evident when the explanations usually contained only one aspect such as "the noun is singular," or "it refers to one thing."

The second problem with exercising the instructional rules is that when the term 'specific' is used (as in Rule 4), the students may not exactly know in what circumstances a given NP would be treated as 'specific'. At first glance the word 'specific' may appear simple, but in terms of article usage, the students do not seem to have a clear understanding as to what exactly they have to look for in order to judge a given reference as 'specific'. It appears that Thai students have not been provided with clear explanations regarding the distinctions between the notions of SR and HK and the importance of considering the notion of HK in determining articles. It might be possible that EFL teachers themselves lack a full understanding of such distinctions. If this is the case, one cannot expect them to be able to teach their students to clearly understand these two notions. This may be a reason why Thai learners have trouble using English articles.

Another problem in exercising these rules is that article usage is context-dependent. Although the students are able to memorize all these instructed rules, application of such rules is not always simple. One cannot simply look for limited structural cues such as a singular/plural form of a noun or the existence of a modifier in determining which article to use. As the low-proficiency students had difficulty in comprehending larger contextual information (due to their limited lexical knowledge), applying article rules—particularly Rule 3 and Rule 4—was their main impediment.

Thus, it can be seen that the NNS-L students relied almost entirely on (S) MLK and they rarely used (P) MLK in referring to referentiality in their choice of *a/an* in indefinite contexts.

(2) Use of *the* in indefinite (Type 3 and Type 4) contexts

As earlier noted, the NNS-L students often formulated non-generalizable rules and collocations when they used *the* in definite contexts. The replacement of *the* in *a/an* contexts in the NNS-L group were also explained based on non-generalizable rules such as:

2.1 the NP refers to a place/a person/an event
2.2 'the' is used with a proper noun
2.3 'the' should be used to begin a sentence

These rule-based non-generalized hypotheses are mainly semantically and syntactically-oriented. In addition to these rules, the NNS-L students also used structural cues such as modifiers to justify their use of the definite article. In the NNS-L group, only a few students used pragmatic (P) MLK (e.g. the NP was 'specific') to explain their choice of *the*.

(3) Use of *Ø* in indefinite (Type 3 and Type 4) contexts

As Table 6.5 indicates, the students in both proficiency groups replaced *a/an* with *Ø* in a similar portion in each context. Among the NNS-L students, the misuse of *Ø* in *a/an* contexts was largely due to the failure to identify head nouns, misjudgment of word classes, and incorrect collocations. A number of students in the NNS-L group used idiosyncratic non-generalizable rules and non-explicit explanations such as impressionistic judgments (I) when they chose the *zero* article (*Ø*).

In summary, the Thai students' problems with using the indefinite article *a/an* varied depending on English proficiency levels. For the NNS-L students, limited vocabulary made it difficult for them to take larger context dynamics into consideration in determining articles. Therefore, they had to rely solely on simple morpho-syntactic or structural cues such as the plural morpheme or the existence of modifiers. Once they considered a given NP to be singular, they chose *a* or *an*; and when it was not, they tended to use *the*. Due to the lack of the ability to consider contexts dynamically, there seems to be a greater likelihood for the NNS-L students to choose *a* or *an* wherever possible. The fact that such learners used *a/an* more than *the* appears to increase the

accuracy rate for their choice of *a/an* in contexts where *a/an* was the target article. For the NNS-H learners, although they began to realize that they had to consider contexts more seriously and dynamically in choosing articles, accurate judgments as to whether references should be identifiable by the hearer turned out to be difficult for them. Explanations offered by the NNS-H students gave some evidence that many of them wrongly judged the references to be part of the assumed hearer's knowledge more than they were supposed to, and thus *the* was overused in a number of items where *a/an* should be required. The difficulty in correctly detecting HK, which was a primary requirement for determining definiteness, was likely a major cause of persistent problems in article usage among Thai learners, even those at a high proficiency level.

6.2.3 Metalinguistic Explanations in the Use of Ø in Indefinite Contexts

As described in Chapter 3, the *zero* article (Ø) marks a plural form of a count noun, and a non-count noun in Type 3 [+SR –HK] first-mention context. It is also used when a referred NP is countable and plural in Type 4 [-SR –HK] non-referentials. In this study, non-count nouns with Ø are classified as Type 3 (first-mention) use and Type 1 (generic) NPs only. As non-count nouns do not have boundaries, references are normally considered to be generic. Except when the reference is specific [+SR], the generally-uncountable entity can then be divisible as in the context of Type 3 (first-mention) use.

Table 6.6 below shows the performances of the three groups of participants in frequency distribution and percentage in Type 3 and Type 4 NP contexts marked with Ø.

Table 6.6: *Frequency Distribution and Percentage of Article Uses in Type 3 and Type 4 Marked with Ø*

Type		Type 3 [+SR -HK]			Type 4 [-SR -HK]		
Target Article		Ø (180 tokens)			Ø (180 tokens)		
Article Choice		Ø	a/an	the	Ø	a/an	the
Group	NS	**111 (61.7%)**	27 (15.0%)	42 (23.3%)	**139 (77.2%)**	2 (1.1%)	39 (21.7%)
	NNS-H	**36 (20.0%)**	54 (30.0%)	90 (50.0%)	**63 (35.0%)**	21 (11.7%)	96 (53.3%)
	NNS-L	**26 (14.4%)**	88 (48.9%)	66 (36.7%)	**32 (17.8%)**	61 (33.9%)	87 (48.3%)

As shown in Table 6.6, Ø uses seemed to be problematic to the participants in all groups, particularly among the Thai students. The NNS students in both proficiency groups produced considerably low rates in the targetlike use of the *zero* article. Even the students with high English proficiency could use Ø correctly with only 20% and 35% accuracy for Type 3 and Type 4 respectively. In the following section, I first summarize the explanations provided by the native English speakers in Ø contexts. I then discuss the reasons given by the Thai learners in relation to their article choices so as to identify problems underlying the use of the *zero* article (Ø) among the Thai learners of both proficiency levels.

6.2.3.1 Metalinguistic Explanations of the Native Speakers in the Use of Ø in Indefinite Contexts

(1) Use of the target article Ø in indefinite (Type 3 and Type 4) contexts

In the NS group, the choices of the target article Ø were explained based on pragmatic (P) and (S) explanations in terms of two main factors:

1.1 a 'non-specific' NP
1.2 the plural form of a given noun

In items where Ø marked a non-count noun, most explanations were based on 'specific reference' of a given NP. The most common reason offered by the native speakers was that a given reference was not specific, or that it referred to an entity in a general sense. The aspect of countability was rarely found in the explanations of the NS students.

In items where Ø marked plural nouns, plurality was mentioned by a number of native speakers in addition to referentiality. Some native speakers incorporated both components in their explanations; for example,

(Quote 62) Ø (ships) – 'ships' is plural and we are not talking about any specific ships

(Quote 63) Ø (suspicious-looking men) – plural, unknown, generalized

(Quote 64) Ø (men) – because 'men' is plural and they were unknown

(Quote 65) Ø (men) – more than one man that has not been explained before

(Quote 66) Ø (detailed descriptions) – plural, non-specific

Not only did the NS students produce the lowest rate of the target use of Ø when compared with the other articles, it was also found that among the native speakers' choices of Ø, non-explicit explanations were given at a high frequency. As one may recall, the quantitative analysis presented in Table 5.9 on page 106 of Chapter 5 revealed that non-explicit explanations (which included impressionistic judgments and no reason provided) constituted a high percentage in the NS choices of Ø in Type 3 and Type 4. As will be seen in the examples below, English articles are often introduced in L1 classes as a small part of a section on 'adjectives.' In this part, only overt forms *a, an*, *the* are described. The following examples illustrate how *a, an*, and *the* are included in textbooks used in schools in the United States.

Grade 3: McGraw-Hill: Language Arts, Texas Edition (2001)

Articles – Rules

"*Articles* are special adjectives. The words *a, an*, and *the* are articles.

Use *a* before a singular noun that begins with a consonant.

a bird

Use *an* before a singular noun that begins with a vowel.

an egg

Use *the* before singular nouns and plural nouns

the nest *the* nests

The articles *a, an*, and *the* can come before singular nouns. Only *the* can come before a plural noun. *The* is used to refer to a specific thing or group.

Jake's family went for *a* ride.

They looked at all *the* sights."

Grade 6: McGraw Hill: Grammar and Composition (2001)

"The word *a, an*, and *the* are special kinds of adjectives. They are called *articles*.

The points to a specific item or items. *A* and *an* refer to any one item of a group. Use *a* before words that begin with a consonant sound. Use *an* before words that begin with a vowel sound.

William Shakespeare is *the* most famous English playwright

Getting a lead role in *a* Shakespeare play is *an* honor."

Grade 10: Language Network: Grammar-Writing-Communication (2001)

Articles

"The most common adjectives are the articles *a, an*, and *the*.

A and *an* are *indefinite articles*. They are used to refer to unspecified members of groups of people, places, things, or ideas.

Use *a* before words beginning with consonant sounds and *an* before words beginning with vowel sounds.

A fan yelled as we looked for *an* exit.

The is the *definite article*, used to refer to a specific person, place, thing, or idea.

The coach yelled as we left through *the* exit."

Grade 12: Prentice-Hall: Writing and Grammar Communication in Action (2001)

Articles

"Three common *adjectives—a, an*, and *the*—are known as *articles*. *A* and *an* are called *indefinite articles* because they refer to any one of a class of nouns. *The* refers to a specific noun and, therefore, is called the *definite article*.

Indefinite: *a* daisy *an* orchid

Definite: *the* stem"

From the above examples, one can see that L1 students are often provided with only a brief summary of article usage in each grade level. These instructions are followed by a few practice exercises. Only the overt articles *a, an*, and *the* are introduced; the *zero* article Ø is not included. One may also notice the word 'specific' used to describe how the definite article *the* can be used. As discussed earlier, metalinguistic knowledge often reflects what has been taught and learned. Since the *zero* article is not included in L1 instruction, it is not surprising if the students often find it difficult to articulate the reasons why Ø is the proper article in a given context.

(2) Use of *a/an* in indefinite (Type 3 and Type 4) contexts

As plural count nouns and non-count nouns do not take the article *a/an*, choices of *a/an* should appear inappropriate. However, the choices of *a/an* were found among the NS students in 15.0% in Type 3 and 1.1% in Type 4. These uses were mainly caused by the failure to detect the plural form of given nouns in some test items. This was

explicitly explained in the students' comments that they considered the referred NPs to be 'singular'. Thus, they chose *a/an* instead of *Ø*.

(3) Use of *the* in indefinite (Type 3 and Type 4) contexts

In the choice of *the*, three main reasons were found in the NS group:

3.1 a 'specific' NP
3.2 a 'generic' interpretation of a given NP
3.3 the plural form of a given noun

The first rationale was consistent with the conventional use of *the* in other contexts where the referred NPs were considered to be part of the assumed hearer's knowledge (often referred to as 'specific' in the native speakers' comments). This rationale was found in the use of *the* in both Type 3 and Type 4 contexts. Although the referred NPs in Type 3 were all first-mention uses, the NS students often associated many references with a previously-mentioned object or event. Such references were accordingly taken for granted to be part of the assumed hearer's knowledge, and thus *the* was chosen instead of *Ø*. For example, the first-mention NPs *broken glass* and *firefighters* were judged as definite because their references were related to the *fire incident* mentioned at the beginning of the paragraph. As earlier discussed, the participants tended to pay great attention to 'specificity' when reading the narrative text. Naturally, one may assume that objects and events in a narrative account are usually connected in some way. Consequently, more references were often wrongly judged to be identifiable by the hearer than they were supposed to, causing *the* to be over-generalized in items in Type 3 first-mention nouns.

Another example of this type can be seen in an item in Type 4 where a few native speakers associated *ships* with the word *sailors* introduced in the preceding sentence, and they commented that:

(Quote 67) the (ships) – specifies as the sailors' ships
(Quote 68) the (ships) – referring to the specific ships of the sailors

Related to the above item, the existence of a modifier was also treated as a device to make a reference definite. Examples below indicate that some native speakers used the time expression *in the old days* to mark definiteness of the noun *ships*.

(Quote 69) the (ships) – the specific ships of *a specific time period*

(Quote 70) the (ships) – 'the old days' specifies 'the ships'

(Quote 71) the (ships) – because 'the old days' is specified, 'the ships' must be as well

The second reason for the use of *the* in Ø contexts was that some native speakers considered a given NP, particularly in Type 4, to refer to 'all entities as a whole' in a generic sense rather than specific/non-specific individual items. Examples of this kind of generic reading can be seen in the following comments which explained why *the* was used in front of *sailors*.

(Quote 72) the (sailors) – *sailors* is plural and *generic*

(Quote 73) the (sailors) – the reference is to all sailors

(Quote 74) the (sailors) – not a specific sailor, but all sailors

Theoretically speaking, in generic use, the plural form of a count noun calls for Ø, not *the*. The definite generic *the*, in contrast, takes a count noun in its singular form, not plural. Thus, it is not consistent with the current linguistic theory when the students used *the* in front of a plural count noun to denote genericity. This casts some doubt as to whether such a structure is acceptable in actual language use. As there were only a small number of explanations which informed this use, it may not be possible to draw any assumptions at this stage. We will investigate article usage in generic contexts in more depth in a later section.

The third rationale used with the choice of *the* was related to (S) MLK. The NS students associated the use of *the* with a plural form of a noun. Explanations such as the following supported the choice of *the* for the nouns *ships, sailors*, *scientists*, and *firefighters*.

(Quote 75) the (ships) – *ships* is plural

(Quote 76) the (sailors) – 'sailors' is plural, so 'a' or 'an' would not fit

(Quote 77) the (scientists) – I can't use 'a' or 'an' for a plural, in this case – '*scientists*'

(Quote 78) the (firefighters) – expresses more than one

(Quote 79) the (firefighters) – '*Firefighter*' is plural, so 'a' or 'an' would not fit

Again, according to linguistic theory, plurality is not a requirement for a noun to take *the*. The definite article *the* (in a non-generic use) can mark a singular or plural form of a count noun as well as a count use of a non-count noun when the reference is identifiable by the hearer. As the students did not include SR or HK in their explicit explanations, it was not clear whether the students implicitly considered a given reference to be 'specific' with regard to the linguistic rule of using *the* or not. Since the analysis in this study was conducted based on the actual statements offered by the participants, inferences should not be drawn on their behalf. Further research may be necessary to investigate to what extent the theory 'a plural noun takes *the*' can be generalized among native speakers in actual language use.

6.2.3.2 Metalinguistic Explanations of the NNS-H Learners in the Use of Ø in Indefinite Contexts

As Table 6.6 indicates, a substantial gap existed in the performance of the NS and NNS students of both proficiency groups in the Ø contexts. The NNS-H group produced only 20.0% and 35.0% targetlike use of Ø in Type 3 and Type 4, whereas the NNS-L students supplied 14.4% and 17.8% accuracy for Type 3 and Type 4 Ø contexts respectively. If one recalls, the differences in the test performances between the two NNS groups in Ø use of Type 3 first-mention NPs (as reported in Table 5.6 on page 97 of Chapter 5) were not different significantly. The following summary of metalinguistic explanations provided by the NNS-H learners for their article choices may help us understand why the use of Ø in this context posed so much difficulty for them.

(1) Use of the target article Ø in indefinite (Type 3 and Type 4) contexts

Although the targetlike choice of Ø yielded considerably low percentages among the NNS-H students, most explanations supporting this article use were found to be somewhat consistent with linguistic theory. Three main reasons provided by the NNS-H students were:

1.1 the plural form of a given noun (or an 'uncountable' noun in the case that a non-count noun was used in two test items in Type 3)

1.2 a 'non-specific' NP

1.3 a first-mention NP

It was found that most explanations for the use of Ø among the NNS-H students were based on the first reason, which was classified as (S) MLK. However, there were some students who combined (S) reasons (in 1.1) with (P) reasons (in 1.2 or 1.3), which suggested that such students had a better understanding of how the *zero* article worked than the majority of the students in this group. Non-generalizable rules were formulated mainly with regard to collocations of Ø with certain prepositions; for example, 'of + Ø' and 'like + Ø'.

(2) Use of *a/an* in indefinite (Type 3 and Type 4) contexts

As shown in Table 6.6, the indefinite article *a/an* was misused more in the first-mention (Type 3) use of Ø than in non-referential (Type 4) contexts (i.e. 30.0% > 11.7%). From the NP structure described earlier, Ø marks plural nouns and non-count nouns; therefore, the use of *a/an*, which marks singular count nouns, is inappropriate. A high proportion of the misuse of *a/an* among the NNS-H students, particularly in Type 3 context, was due to two main reasons:

2.1 the failure to detect a plural form of a given noun

2.2 the misjudgment of noun countability.

From the students' explanations, some NPs appeared to be more problematic in the detection of countability and their plural form. For example, the irregular plural form of the noun *men* was not recognized by many students, and thus, it was wrongly treated as a singular noun for the students who used *a* for this item.

The plural noun *people* was found to be the least problematic. Twenty-four NNS-H students (equivalent to 80%) chose the correct article. This was probably due to the students' familiarity with the word *people* commonly used without an overt article to refer to *people* in general.

Nouns that begin with a vowel sound, such as *echoes* and *echo-sounders*, were likely to draw the students' attention to the rule of the consonant/vowel distinction with respect to the sound of the following word rather than other important components. This phono-syntactic rule seemed to be so influential that it often caused the students to overlook other elements (e.g. plurality) that were necessary for determining a proper article.

In the detection of noun countability, a non-count noun whose countability was mis-detected by many students was *glass*, as in the NP *broken glass.* In this item, none of the students in the NNS-H group chose the target article. The students' explanations indicated that a large number of students misread the meaning of the polysemous word *glass*; they interpreted it as a count noun meaning a container used for drinking. Thus, nineteen NNS-H students chose *a* instead of *Ø*.

While the non-count use of *glass* was misjudged by many NNS-H students, the word *blood* was found to be correctly detected as a non-count noun in its given context by a majority of students. In this item more than half of the NNS-H students chose *Ø,* and most of them justified their choice by saying that *blood* was an uncountable noun. Among those who chose *Ø*, two students incorporated referentiality with countability and explicitly stated that *blood* was 'uncountable' and 'non-specific'. In fact, for a non-count noun to take *Ø*, its reference must be unidentifiable by the hearer [-HK]; otherwise *the* is possible. It may be assumed that the students' correct choice of *Ø* for *blood* without considering its [-HK] reference might, again, be due to the students' familiarity with the typical use of *blood* in a general sense without an article. When *blood* needs to be identified as to whose *blood* it belongs to, a possessive adjective is normally used; for example, *my blood* or *his blood*, rather than the definite article.

(3) Use of *the* in indefinite (Type 3 and Type 4) contexts

Like the native speakers, the NNS-H students' variable choice of *the* in *Ø* contexts was based on three main factors.

3.1 a 'specific' NP

3.2 the plural form of a given noun

3.3 a 'generic' interpretation of a given NP

In judging references to be 'specific', many NNS-H students tended to rely on structural cues, such as the existence of modifiers, as a convenient device that marked definiteness. Thus, whenever there was a modifier attached to the reference, these students often automatically chose *the*. Examples that serve to illustrate this can be found in the noun phrases: *broken glass, detailed descriptions,* and *suspicious-looking men*. In these NPs, the preposed adjectives *broken, detailed*, and *suspicious-looking* were treated as an indicator of 'specificity' by the students, resulting in the insertion of the definite article *the*.

(Quote 80) the (broken glass) – specifies the glass that was broken

(Quote 81) the (detailed descriptions) – 'descriptions' is specified by 'detailed'

(Quote 82) the (suspicious-looking men) – 'men' is identified by 'suspicious-looking'

Similarly, the prepositional phrase *in the old days* was treated as if it served to identify the 'specific' *ships* of a 'specific' time period, and thus *the* was used instead of *Ø*.

In addition to the existence of modifiers, references were often taken for granted that they were a part of the assumed hearer's knowledge when the students brought too much extra-linguistic knowledge and their own presuppositions when making article choices. As such, the students considered first-mention nouns such as *firefighters* to be 'specific' based on the assumption that it referred to the 'specific' firefighters who were present at the fire incident mentioned in the preceding sentence. By the same token, as the story described *the dolphin*, the first-mention noun *scientists* was assumed to refer specifically to the scientists who were researching the life of the dolphin. Thus, *scientists* was judged to be 'specific', and *the* was chosen instead of *Ø*.

The second reason for the choice of *the* among the NNS-H students was due to the theory that a plural noun takes *the*. This non-generalizable hypothesis may have been influenced by a basic instructional rule that when a noun is countable and singular, *a/an* is used; thus, when it is plural, *the* should be used. It was found that this non-generalizable theory was commonly observed among a number of the Thai learners. As

one may recall, this rule was also observed by many native English speakers when they used *the* to mark plural nouns.

The use of *the* was also rationalized in terms of a generic reading, particularly with NPs in Type 4 non-referential use such as *sailors, scientists*, and *people.* Many students considered references of such nouns to refer to all entities as a whole group. This was also consistent with explanations offered by some native speakers as discussed in the previous section. For example,

(Quote 83) the (sailors) – referring to all sailors as a whole
(Quote 84) the (scientists) – referring to all scientists as a whole

To sum up, one can see that accurately using Ø requires a proper detection of a number of components; namely, the referentiality features SR and HK, countability, and number. Number did not appear to pose a great problem to learners at a high proficiency level; however, noun countability could more easily be misjudged, particularly when the students misread the intended meaning of a word. It is notable that in the contexts where Ø was required, the NNS-H learners tended to misuse *the* to a great extent (50.0% in Type 3 and 53.3% in Type 4). This appeared to be due to the students' excessive introduction of external knowledge (such as their own presupposed beliefs) into their readings. References were often taken for granted to be part of the assumed hearer's knowledge, and thus *the* was overused in items in Type 3. In addition, references were also wrongly judged as 'specific references' based on the existence of modifiers.

As Ø stacks multiple complicated functions, the NNS-H students appeared to show apprehension at the use of Ø (or nonuse of an article). So, when they were not confident about using Ø, they often tried to find a justification for choosing *a/an* or *the* instead. As far as students' explanations indicated, the choice of Ø most frequently yielded non-explicit reasons in the NNS-H group (as well as the NS group). This explains why learners, even at a high proficiency level, find the *zero* article difficult to use and its usage difficult to explain.

6.2.3.3 Metalinguistic Explanations of the NNS-L Learners in the Use of Ø in Indefinite Contexts

(1) Use of Ø in indefinite (Type 3 and Type 4) contexts

As demonstrated in Table 6.6, the NNS-L students had only 14.4% and 17.8% targetlike use of the *zero* article in Type 3 and Type 4 contexts respectively. One common explanation that was consistent with linguistic theory found among the NNS-L students was that the given NP was plural. Other reasons mainly contained non-generalizable rules or were based on non-explicit explanations, such as guessing. Referentiality features were not mentioned by the students in the NNS-L group. In some items, the target article Ø was chosen, but for linguistically-inappropriate reasons. For example, a number of students merely decided to use or not to use an article based on their judgment of the word that followed. An article would be inserted when the students regarded the following word as a noun. Thus, if they failed to recognize a noun, they would assume that no article should be inserted. Based on this hypothesis, the students chose to put Ø in front of *broken glass* and *suspicious-looking men*, and explained that an article was not needed because '*broken*' and '*suspicious-looking*' were not nouns.

(2) Use of *a/an* in indefinite (Type 3 and Type 4) contexts

The non-targetlike use of *a/an* in Ø contexts constituted about half of the responses in Type 3 (48.9%) and about one-third of the responses in Type 4 (33.9%). The students' choice of *a/an* instead of Ø was found to result from two main semantico-syntactic (S) reasons:

2.1 failure to detect countability
2.2 failure to detect the plural form of a given noun

A non-count use of nouns such as *glass* and *blood* appeared to be problematic for the NNS-L students. As far as the students' explanations indicated, many NNS-L students merely determined the use of *a/an* based on whether a given noun had the plural morpheme *–s* or not. If there was no *–s*, the students would perceive it as a singular noun, and thus *a/an* was chosen to be used. As countability is a more complex aspect,

the NNS-L students based their article choices on this feature less frequently than on the distinction between the singular and plural form of a noun.

In the NNS-L group, failure to recognize irregular forms of plural nouns such as *men* and *people* was rather common, for most NNS-L students often judged plurality simply by looking at the existence of the inflectional morpheme *–s*. However, it was found that more than half of the NNS-L students still failed to recognize the plural form of nouns with *–s* that start with a vowel, such as *echoes* and *echo-sounders*. In these two items, *an* was used instead of *Ø* by a majority of the NNS-L students. It appeared that the phono-syntactic rule regarding the contrast between *a* and *an* was so influential that the students often automatically inserted *an* once they saw a vowel at the beginning of the next word, while ignoring other important components (e.g. plurality).

While the problem in using *a* as opposed to *an* was rare in the NNS-H group, it was found that a number of NNS-L students still lacked a proper understanding in the use of *an*. For instance, some NNS-L students explained that they used *an* to connect words or phrases, which suggested that such students confused the article *an* with the conjunction *and*. Due to the lack of understanding of the use of this article, many non-generalizable rules were formulated; for example,

a = singular, thus *an* = plural; *a* = count nouns, thus *an* = non-count nouns

(3) Use of *the* in indefinite (Type 3 and Type 4) contexts

The inappropriate choice of *the* among the NNS-L students comprised about one-third of the responses in Type 3 (36.7%) and about half of the responses in Type 4 (48.3%). A frequently-used explanation for choosing *the* was a non-generalizable hypothesis that a plural noun should take *the*. There were only a few students who referred to referentiality features to support their choice. Explanations that were inconsistent with linguistic theory were found in a greater number in the students' choice of *the* than in their use of *a/an*. These included non-generalizable rules such as:

a = singular, thus *the* = plural; *a* = count nouns, thus *the* = non-count nouns
a = common nouns, *the* = proper nouns; *a* = action, *the* = event or situation
a = animal or thing, *the* = place or person

a = object position, *the* = subject position

According to the hypotheses above, the NNS-L students often relied on rule-based hypotheses in making article choices. Of the three articles: *a/an, the*, and, *Ø*, the NNS-L students appeared to be more capable of providing explicit reasons when they chose *a/an* or *the* than when they chose to use *Ø*. For the choice of *a/an*, the students mainly relied on some basic instructional rules in their explanations. When the conditions for using *a* (or *an*) did not pertain, the students formed a variety of rule-based non-generalizable hypotheses from their own learning experience to accommodate their choice of *the*. This was likely due to their limited ability to consider larger dynamic contexts for determining the use of *the*. In the students' use of *Ø*, however, there seemed to be a gap in their mental representations (rules). The students, particularly those in the NNS-L group, had more difficulty providing explicit reasons when they chose *Ø*. This problem may be due to the fact that article instruction mainly focuses on the overt articles *a/an* and *the*. The *zero* article (*Ø*) is often left unexplained. It is frequently taken for granted that *Ø* (no insertion of an article) should be applied when the conditions for *a/an* or *the* do not pertain. For low-proficiency students, though, this assumption does not seem to work. As these students are less proficient, they may need more explicit instruction as to when and why an overt article is not needed. The lack of a clear understanding of the proper use of *Ø* (or simply put, a nonuse of an overt article) might be a reason why the NNS-L students in this study underused *Ø* to a large extent. From investigating the reasons for the use of *Ø*, non-explicit explanations were found at a high rate. Statements offered by the students included: "I'm not sure," "I just guessed," "I think no article is needed" and "it sounds better without an article." Explicit reasons for the use of *Ø* were based mainly on non-generalizable hypotheses and misjudgments of words. There were only a small number of students who used *Ø* correctly with an explanation that contained a partially accurate description with regard to linguistic theory.

In summary, the NNS-L learners relied mainly on a few basic rules from which they formulated more idiosyncratic non-generalizable mental representations about the overt articles *a/an* and *the*. Due to limited lexical knowledge, these learners had to rely on static local structural cues, such as the plural morpheme *–s* in applying these rules in given contexts. Rules for using *Ø* were rarely cited by the NNS-L students. Moreover,

Ø was used in a small percentage. This was evident that the students did not have a clear understanding for the use of Ø. As one can see, the accurate use of Ø did not yield a significant difference between the two proficiency groups. One may conclude that, as instruction on Ø is often lacking in EFL article lessons, usage of Ø is the most problematic for Thai students even when the proficiency level increases.

6.2.4 Metalinguistic Explanations in the Use of Generic Articles

Generic reference is used to denote a "class of entities" as a whole rather than a specific or non-specific individual member of the class. Generic descriptions can take all three kinds of articles in the following structures (Please refer to Table 3.1 in Chapter 3):

1. Generic-*the*: [the [+count] [+sing]] NP
2. Generic-*a/an*: [a/an [+count] [+sing]] NP
3. Generic-Ø: [Ø [+count] [-sing]] NP or [Ø [-count]] NP

As mentioned in the literature review, there are pragmatic and semantic differences among the three kinds of articles used in generic descriptions. The choice of articles seems to depend on discourse considerations and the predicates which denote the type of proposition regarding the generality implied in a given generic NP. In Huebner's model, the generic use is characterized as Type 1 [-SR, +HK]. Table 6.7 below displays the number of tokens of the targetlike and non-targetlike article choices in Type 1 generic contexts, with percentages in parentheses.

Table 6.7: *Frequency Distribution and Percentage of Article Uses in Type 1 Generics*

Type 1 [-SR +HK]									
Target Article	Ø (120 tokens)			a/an (120 tokens)			the (120 tokens)		
Article Choice	Ø	a/an	the	Ø	a/an	the	Ø	a/an	the
NS	**92 (76.7%)**	1 (0.8%)	27 (22.5%)	0 (0%)	**65 (54.2%)**	55 (45.8%)	0 (0%)	9 (7.5%)	**111 (92.5%)**
NNS-H	**36 (30.0%)**	20 (16.7%)	64 (53.3%)	7 (5.8%)	**68 (56.7%)**	45 (37.5%)	5 (4.2%)	60 (50.0%)	**55 (45.8%)**
NNS-L	**22 (18.3%)**	58 (48.3%)	40 (33.3%)	7 (5.8%)	**78 (65.0%)**	35 (29.2%)	4 (3.3%)	67 (55.8%)	**49 (40.8%)**

As one may recall, the generic use of articles is a fuzzy NP type. The indefinite generic-*a/an*, in particular, yielded no statistically-significant differences in test

performances among the participants in the three sample groups. In the three sections that follow, I first summarize the metalinguistic explanations provided by the native speakers with regard to their performance in the generic context. Then I report the explanations offered by the Thai students in NNS-H group before I compare them further with those given by the NNS-L students.

6.2.4.1 Metalinguistic Explanations of the Native Speakers in the Use of Generic Articles

As shown in Table 6.7, the definite generic-*the* seemed to be least problematic among the native English speakers. In generic-*the* use, the variable choice of *a/an* was found in only 9 tokens (7.5%), whereas in the generic-*a/an* use, the choice of *the* was found in 55 tokens, or 45.8% of the total responses. From the explanations offered by the NS participants, the choices of *a/an* and *the* largely depended upon the individual's perception of generality of the referent whether it is interpreted as a single member representing the whole class or whether it represents the whole class itself. As the definite generic-*the* is the most formal way to express generality with the overall meaning of "total or whole," the test items that called for *the* did not appear to pose difficulty among the NS participants. Most native speakers explained their choice of *the*, in *the dolphin, the lion, the lioness*, by saying that it identified a specific animal species.

Generic-*a/an*, on the other hand, requires a more subtle interpretation as it presupposes that the speaker is referring to one entity as a representative of the whole class. As such, one needs to carefully consider whether the context is talking about an example of one entity which represents all entities, or whether it is referring to all entities as a whole. For example, in the sentence below, the pronoun *it* lends its singularity to the reference of *dolphin* toward representing a single member of the whole dolphin species.

'By sending out sounds which bounce back like echoes, (16) _a_ dolphin is able to tell whether a rock or fish lies ahead of *it*.'

In item 16, there were some NS participants who explicitly explained their choices of *a* for *dolphin* as follows:

(Quote 85) a (dolphin) – general description of members of species

(Quote 86) a (dolphin) – this sentence is giving an example of one dolphin sending out a signal

(Quote 87) a (dolphin) – referring to just one dolphin which represents all dolphins

Most NS students who chose *the* for this item considered the reference in terms of the whole group, while two native speakers judged the reference to be 'specific' instead of 'generic'.

Similarly, in the following two sentences, the use of *his* in the second sentence suggests that a single entity is being compared.

'Unlike most cat species, (40)_ *a* _male lion is different from (41)_ *a* _lioness. *His* head is much larger than that of (42)_ *a* _female lion.'

As 'his head' in the second sentence refers to the head of 'a male lion' in the preceding sentence, the comparison between items (40) and (41) should then be between a member of the male and female lion species. This is also the case with item (42). In these three test items, some native speakers explained their choice of *a* as follows:

(Quote 88) a (male lion) – a general, single member of the group

(Quote 89) a (lioness) – talking about the differences of a male lion and a female

(Quote 90) a (female lion) – using particular words like describing one (=his)

(Quote 91) a (female lion) – because we are talking about a single lion; we will compare his head to a single female

(Quote 92) a (female lion) – 'head' referred to a 'his' earlier in sentence, meaning we are comparing 'one head'

The NS students who chose *the* for these items explained that the given NP referred to a specific gender of lions—i.e. a certain group of male and female lions.

It should be noted, however, that in these three test items, there were a relatively large number of changes in responses between *a* and *the*. This suggested that although a generic interpretation was recognized, the NS students were indecisive in their article choices for these three NPs. Only a few native speakers explicitly stated that they chose *a* based on a syntactic cue (such as 'his') which indicated that a member of a whole

species was being referred to. Many native speakers offered a variety of reasons to support their choices of *a*; for example,

(Quote 93) a (male lion) – one of many

(Quote 94) a (male lion) – any male lion

(Quote 95) a (lioness) 'lioness' is general

(Quote 96) a (lioness) – non-specific single lioness

(Quote 97) a (female lion) - generalization

(Quote 98) a (female lion) – a lion is not being specified

As one can see, a categorical reading of a definite generic (generic-*the*) use is easier to identify from the discourse than an indefinite generic (generic-*a/an*) use. Since all generic readings denote the generality of entities in some way, the native speakers were likely pre-disposed toward a categorical interpretation, and thus *the* appeared to be the preferred article for expressing genericity.

For the generic use of the *zero* article, the target article was chosen in 92 tokens (76.7%) by the NS students; *a* was variably chosen in only 1 token (0.8%) and *the* in 27 tokens (22.5%). The choice of the target article *Ø* was explained mainly based on the generic interpretations of the given NPs. The terms frequently used among the native speakers were "general, non-specific, not specified, and generalized." In items where plural nouns were used, a few native speakers combined the plural aspect with the generality of the referred NP, while a few others associated their use of *Ø* solely with a plural form of given nouns.

The use of '*the* + plural nouns' were also observed among some native speakers to denote generic interpretation. The following examples are explanations given by the native speakers who explained their use of '*the* + a plural noun' based on their perception that it represented 'a specific group or species as a whole'.

(Quote 99) the (dolphins) – it is talking about dolphins as a whole species – 'the dolphins'

(Quote 100) the (dolphins) – refers to a specific animal group

(Quote 101) the (dolphins) – This is referring to all dolphins as one group

(Quote 102) the (lions) – points out lions as a group

(Quote 103) the (lions) – referring to a specific species

(Quote 104) the (lions) – describes the group of lions

Theoretically speaking, the definite article *the*, which denotes a categorical interpretation (i.e. as a whole group/species), marks a singular count noun. However, as six native English speakers in this study (equivalent to 20%) indicated that they used '*the* + a plural count noun' to denote a generic reading, it is doubtful whether this use is acceptable in current English language use. This issue may require further investigation with a larger sample group of native English speakers.

As noted in the literature review chapter, generic article usage has received minimal coverage in ESL/EFL curricula. Since generic interpretations are based entirely on context, Master (1988) suggests that the detailed description of the generic article should best be taught after all other aspects of the article system have been thoroughly practiced. This suggestion is supported by the result of the metalinguistic explanations offered by the Thai students in the present study. An examination of the accuracy rate of Type 1 (generic) items offered by the Thai students alone may not reflect a true understanding of the important aspects of the generic use of articles, particularly among the students with low English proficiency. The following is a summary of explanations supplied by the NNS-H students for their article uses in Type 1 (generic) NPs. I will then summarize the explanations of article uses in Type 1 generic contexts given by the NNS-L group thereafter.

6.2.4.2 Metalinguistic Explanations of the NNS-H Learners in the Use of Generic Articles

As one may recall, the quantitative analysis revealed that there was no statistically significant difference in performance between the NNS students in the two proficiency groups in all three forms of generic articles. The explanations provided by the students in the NNS-H group seemed to indicate that they lacked a clear understanding of the semantic difference between generic-*a/an* and generic-*the* uses. Although many students expressed their recognition of the generic reading of the context, they tended to think that the articles *a/an* and *the* could be used interchangeably when a given NP was a singular count noun. One notable point was that *a/an* was found

to be used more frequently than *the* when a generic singular NP in question was introduced for the first time in the text, whereas *the* was chosen more often with previously-mentioned nouns. The students' decision about *a/an* as opposed to *the* was not made on the basis of a true understanding of the semantic difference between generic-*a/an* and generic-*the* as denoting a single example representing a class as opposed to a categorical interpretation. This suggests that the NNS-H students tended to place great attention on referentiality, particularly when using *the*. The use of '*the* + a singular count noun' was strongly connected to Type 2 article usage (i.e. the notion of a 'specific reference') rather than a 'generic reference'. Only a small number of students expressed an understanding that a generic NP with *the* denotes a specific group or species of things, people or animals. Yet, these students did not clearly explain how they interpreted the generic context in which they decided to use *a/an* as opposed to *the*.

With regard to the generic context of a non-count noun, the target article Ø was chosen by only five students (16.67%). As far as their explanations indicated, a majority expressed their recognition of the generic concept of '*man*' in its given context, as many of them explained that it was used to refer to 'any' man in general. However, as *man* was regarded by the students as a singular count noun, they chose to use *a* instead of Ø. The variable choice of *the* in this context by a small number of students was explained based on their recognition of the generic reading and their understanding that '*the* + a singular count noun' denotes a generic interpretation. Such students indicated that they interpreted *man* as referring to the whole human race; no one referred to it as denoting a member of the human species. As *man* was mentioned for the first time in the very first sentence, it was not logical for the students in the NNS-H group to connect this use with 'specific reference' in Type 2 article usage although they used *the*.

Interestingly, however, the students who happened to choose the target article Ø in front of the non-count use of *man* actually explained their choice based on the wrong assumption that 'Ø + a singular count noun' denoted a generic reading. As they considered *man* to be a singular count noun in a generic context, they chose Ø due to this notion.

For plural countable generic nouns, *the* was used more than the target article Ø among the students in the NNS-H group. One of the reasons for the students' choice of

the was, again, that the given NP had already been mentioned. Some students used *the* simply because they thought plural nouns required *the*, not Ø (as discussed in an earlier section). It was also found that some students thought that the combination of '*the* + a plural count noun' denoted a generic reading. Only a small number of students chose the target article Ø for a linguistically appropriate reason. These students had the proper understanding that 'Ø + a plural count noun' indicated a common generic usage.

To sum up, as their proficiency developed, the NNS-H students began to realize that they had to depend on context to determine article choices. Accordingly, the recognition of generic article use, which is context-dependent, seemed to emerge among a majority of students in the NNS-H group. However, they still appeared to lack a clear understanding of the distinction between the uses of generic-*a/an* and generic-*the*. The NNS-H students tended to use the combination of '*a/an* + a singular count noun' to denote generality, which was somehow connected to the notion of Type 4 non-referential use. As supported by explanations offered by many native speakers in this study, the distinction between the generic-*a/an* use and non-referential (non-specific indefinite) use can be very difficult to define. One can see from several examples provided in the previous sections that the NS and NNS-H students did not clearly distinguish between generic and non-referential interpretations.

The use of *the* by the NNS-H students indicated a more formal way to express genericity in terms of the whole group/class representation. However, only a small number of students understood this usage. Most students treated the definite article as if it was being used as a Type 2 context rather than being used as a generic article.

The targetlike use rate for generic-Ø was very low among the NNS-H students. This is rather uncommon, given that the students usually learn a basic rule that a plural noun with no article can make a reference generalized. Yet, the students appeared to show apprehension at a lack of an article, and thus *the* was often inserted, particularly when a given reference was wrongly considered to be anaphoric usage.

Since generic articles seem to receive minimal interest in pedagogy (Master, 1988), generic usage of articles is still problematic for Thai students even at a high-proficiency level.

6.2.4.3 Metalinguistic Explanations of the NNS-L Learners in the Use of Generic Articles

As noted above, the targetlike use of Type 1 generic items in the NNS-L group was not significantly different from that performed by the NNS-H students. However, if one calculated accuracy based on the responses with appropriate reasons that reflected generic article usage, the performance of Type 1 articles among the NNS-L students would drop tremendously. It was found that the NNS-L students often failed to recognize the context as generic. They chose articles based mainly on countability and number of a given noun, without incorporating a generic reading of the context. Thus, when a reference was a singular count noun (which appears in generic-*a/an* and generic-*the* contexts), the NNS-L students tended to choose *a/an* almost automatically. This probably explains why these students had a higher rate of targetlike use in the generic-*a/an* context than that of the NNS-H students (and also the NS participants). As for the choice of *the* among the NNS-L students, it was found that their explanations were based on numerous reasons, most of which were inconsistent with linguistic theory.

In generic-Ø contexts, the NNS-L students used the target article Ø in only a small proportion. Their reasons did not reveal the recognition of generic references, but were based on their consideration that a given NP was plural. Many responses were based on non-generalizable hypotheses or non-explicit reasons, such as guessing.

As discussed earlier, generic readings rely heavily upon the context. Thus, the recognition of generic references seems to emerge at a later stage in the interlanguage, possibly when students develop higher comprehension ability. The NNS-L students in the study failed to treat the generic use of articles as generic due to their limited ability to comprehend the given context. Although the students chose the correct articles, their explanations indicated that they did not have a linguistically appropriate understanding of the use of articles in given situations.

6.3 Summary

This chapter presented metalinguistic explanations of article usage employed by native English speakers and Thai learners of high and low proficiency levels. It attempted to identify the problems underlying English article use in terms of

definiteness, genericity, and countability in the interlanguage development among Thai learners.

First, quantitative analyses were conducted on the participants' targetlike and non-targetlike choices of each article. It was found that the NNS-H students, like the native speakers, tended to overuse the definite article *the*, whereas the students in the NNS-L group used the indefinite article *a/an* more than necessary. The *zero* article was underused in all contexts by all groups of participants. In investigating in greater depth how articles were distributed across NP types, the accuracy rate for using *the* in Type 2 was highest in the NS and the NNS-H groups. The NNS-L students performed better with the indefinite article *a/an* in Type 3 and Type 4 than the definite article *the* in Type 2. In Type 1 generic use, which is a fuzzy NP type, there was no statistically significant difference in performance between the NNS-H and NNS-L groups in all three nominal types of article.

Based on the quantitative analyses, the participants' metalinguistic explanations for their article uses have provided useful information regarding metalinguistic knowledge (MLK) of article usage in definite contexts, indefinite contexts, and generic contexts. In the NNS-H group, the students considered contexts more seriously and dynamically. They selected articles based primarily on pragmatic (P) MLK. Thus, they were more accurate in using the definite article in definite contexts than the students in the NNS-L group. With limited lexical knowledge, the NNS-L students had to rely on local structural cues in determining articles. In so doing, they employed a few basic rules regarding article usage and often exercised each rule independently. As the instructional rules for using *a/an* requires syntactic considerations, the NNS-L students initially chose *a/an* once they thought that a given noun or noun phrase called for the indefinite article. The fact that they often used semantico-syntactically (S) MLK in determining articles might explain why they overused *a/an* in almost all environments.

With their greater proficiency, the NNS-H students began to realize that the definite article *the* was deeply context-oriented. However, they had difficulty in judging what should have been considered identifiable by the hearer in order to use *the* appropriately. What was occurring in this instance was that they often wrongly judged

numerous references to be part of the assumed hearer's knowledge [+HK], and thus they overused *the* to a certain extent.

The *zero* article (omission of an article) was problematic for the students in both proficiency groups. Most article rules provided in EFL classes often give descriptions for the usage of overt articles *a/an* and *the*; no clear descriptions have always been provided for correctly using the *zero* article (Ø). Thus, not only did the NNS students in both groups feel uncertain of not using an overt article, but they also had difficulty providing an explicit reason when they finally decided to use Ø.

The performance of the NNS-L students with respect to the generic use of articles (Type 1) did not reveal a significant difference compared to the performance of the NNS-H students. However, when investigating the explanations offered by the students in the two NNS groups, it was found that most students in the NNS-H group recognized generic readings, but they did not have a clear understanding of the distinction between generic-*a/an* and generic-*the*. The NNS-L students, in contrast, did not recognize generic references in the given contexts. Their metalinguistic explanations indicated that they chose articles based solely on countability, number and other limited structural cues. This suggests that although the performances of the NNS-H and NNS-L students did not reveal a statistically significant difference, one can conclude from the MLK analysis that the accuracy rate of the NNS-L students did not completely reflect their understanding of the generic use of articles.

In the next chapter, I will conclude the main findings of this study, which lead to a few important implications about the aspects of English articles for which EFL instruction can be made more effective.

Table 6.8: *Explanations Inconsistent with Linguistic Theory*

Classification	Examples
Non-generalizable collocation rules	1. Preposition + article (e.g. "at + the," "on + Ø," "on + the," "in + Ø," "of + a," "of + Ø") 2. Article + noun (e.g. "the + hotel," "Ø + manager," "Ø + prey" 3. Adjective + article (e.g. "the + large" 4. Connector + article (e.g. "that + Ø," "wh- + Ø"
Non-generalizable hypotheses	1. The sentence should start with 'the' 2. The same article should not be used successively in one Sentence 3. Plural forms of count nouns take 'the' 4. Non-count or mass nouns take 'the' 5. Non-count or mass nouns take 'a' 6. If a noun is modified by an adjective, it has to take 'the' 7. The subject of a sentence should take 'the' 8. The object of a sentence should take 'a' 9. 'the' is used at the end of the sentence 10. 'the' precedes nouns referring to places or objects
Misreading of texts	1. Syntactic misreading (e.g. the student failed to identify head-nouns) 2. Lexical misreading (e.g. the student misread a word, misjudged a word class or misunderstood the meaning of a word)
Noun countability and number	1. Noun countability (the student failed to detect the countability of nouns) 2. Number (the student failed to detect a singular or plural form of a noun)

Classification	Examples
Referentiality	1. Mis-detection of 'specific reference' (SR) 2. Mis-detection of 'hearer's knowledge' (HK) 3. Mis-detection of both 4. Failure to consider referentiality
Mis-detection of vowel / consonant distinction	1. Failure to consider the vowel/consonant distinction of the beginning sound of the word next to the article 2. Failure to understand the distinction between 'a' and 'an' 3. Mistook 'an' for a connector
Misunderstanding of generic uses	1. "Ø + singular count nouns" = generics 2. "the + plural count nouns" = generics 3. "a + generics" = "the + generics" 4. Failure to recognize generic readings 5. Other misunderstandings of generic uses of articles

CHAPTER VII
CONCLUSIONS, IMPLICATIONS AND RECOMMENDATIONS

The purpose of this study was to examine the metalinguistic knowledge of the English article system employed by Thai EFL learners and compare it with the metalinguistic reasoning offered by native English speakers when making article choices in given situations structured under a theoretical framework. The goal was to understand whether or not there were any differences in the metalinguistic knowledge of native speakers and Thai learners that would yield useful information regarding article acquisition problems among Thai learners. The study also aimed to find out whether there were significant discrepancies between the rules or explanations that learners based their judgments upon and those that were consistent with linguistic theory.

Performance-based analysis alone does not provide an accurate picture of what learners know and what they do not know about English article use. It may even distort our understanding of learners' knowledge and actual abilities in article usage when they choose a correct article but for a linguistically inappropriate reason. Therefore, in this study, two types of metalinguistic measures were used: a fill-in-the-article test and a metalinguistic knowledge elicitation questionnaire. With the combination of the two tasks, one is more able to verify the participants' performance on the fill-in-the-article test, and understand what kind of metalinguistic knowledge they used in choosing articles in given contexts. The ultimate goal of the present study is to identify where learners' article problems originate. It is hoped that the information gained from this study will help teachers find a more effective way to remedy these problems.

In light of the above objectives, the present study sought to answer the following questions addressed in Chapter 1:

(1) What is the metalinguistic knowledge used by native English speakers and Thai EFL learners of high and low proficiency levels with regard to the English article system?

(2) What are the similarities and differences in the metalinguistic knowledge formulated and used in determining English articles by native English speakers and Thai learners of high and low proficiency levels?

(3) What are the problems underlying the use of English articles for Thai learners with regard to the metalinguistic knowledge of definiteness, genericity, and countability?

The following hypotheses were therefore formulated and tested:

Hypothesis 1:

Native English speakers' metalinguistic knowledge of the English article system is relatively stable and pragmatically-oriented. Thai learners' metalinguistic knowledge is primarily syntactically and semantically-oriented. However, the metalinguistic knowledge of Thai learners with high proficiency is more pragmatically-oriented than learners with low proficiency levels.

Hypothesis 2:

The metalinguistic knowledge of the English article system formulated and used by native English speakers and Thai learners is similar syntactically and semantically, but different pragmatically.

Hypothesis 3:

The problems underlying the use of English articles for Thai learners are mostly pragmatically-oriented. Problems also arise from the surface syntactic structures of noun phrases in English.

In the following three sections in this concluding chapter, I will first discuss the main findings of the present study which tested the hypotheses described above. In the second section, I will suggest implications drawn from the study. The last section will offer some recommendations for further research.

7.1 The Main Findings of the Study

In terms of performance on the fill-in-the-article test, the scores on the test increased according to the proficiency groups although there was a sizable gap in performance between the native English speakers and the high-proficiency students in this study. The analysis of test performance based on Huebner's pragmatic-semantic NP type revealed that the scores for Type 2 [+SR, +HK] items were the highest item-type scores across the groups. However, if we analyze the scores of each NP type by article, the performance on the use of *the* in Type 2 was highest among the native speakers (96.4%) and the high-proficiency students (73.1%), whereas in the low-proficiency group the scores for using *a/an* in Type 3 [+SR, -HK] and Type 4 [-SR, -HK] were higher than the performance on *the* (61.1%, 60.0% > 53.9%). The students in both proficiency groups had the most trouble with the *zero* article. With regard to Type 1 [-SR, +HK] generic articles, generic-*a/an* appeared to be problematic for all groups of participants in the study. The low-proficiency students outperformed the high-proficiency students and the native English speakers (65.0% > 56.7% > 54.2%).

As this study aimed to investigate the metalinguistic knowledge (MLK) used by the participants in determining English articles in the domain of written language, the metalinguistic explanatory task was the central measurement of the study. The results of the metalinguistic explanations were used to determine the kinds of MLK employed by the Thai EFL learners of high- and low-proficiency levels as opposed to the native English speakers with regard to article usage. The results were also used to compare and contrast the similarities and differences in the MLK formulated and used among the three participant groups. Such findings answered the first two research questions and tested the corresponding hypotheses.

The metalinguistic explanations in the participants' targetlike use of article were categorized into 2 main categories: explicit reasons and non-explicit reasons (Please refer to Section 4.5 in Chapter 4). The former was subcategorized into: pragmatic (P), semantico-syntactic (S), and the combination (PS). The latter was divided into: impressionistic (I), guessing/pseudo-guessing (G), and no response (N). The results of the metalinguistic analyses revealed that the overall MLK used by the high-proficiency students in determining articles in the fill-in-the-article test was more pragmatically-

oriented and more consistent with the native speakers' norms than the low-proficiency students.

In terms of NP types, Type 2 and Type 1 articles were explained based on (P) MLK most frequently by the NNS-H students. As Type 3 and Type 4 articles (*a/an* and *Ø*) incorporate countability and number, in addition to referentiality features, (S) MLK was used more than (P) MLK in the choices of these articles among the NNS-H students. The results were in line with those in the NS group, whereby (P) MLK was used primarily with Type 2 and Type 1 NPs. However, the NS participants also explained their article choices in Type 4 using (P) MLK more frequently than (S), though the use of (S) and (PS) MLK was also high. In the NNS-L group, the results of the MLK analyses were different from those of the NS and NNS-H groups. The NNS-L students explained their article choices based almost entirely on (S) MLK, and they used (P) and (PS) MLK in very small percentages.

When examining the MLK used in determining the article that marked each NP type (referred to as *NP sub-type*), the results showed that the NNS-H students used (P) MLK with *the* in Type 1 and Type 2 more frequently than other MLK sub-categories. This was consistent with the results of the native speakers. In the use of *a/an*, the results were somewhat different between the NS and the NNS-H groups. While the native speakers used (P) MLK more frequently with *a/an* in Type 1 and (S) more frequently with *a/an* in Type 3 and Type 4, the NNS-H students used (P) MLK more frequently with *a/an* in Type 3 and (S) more frequently with *a/an* in Type 1 and Type 4. However, it was notable that in the use of *a/an* in both groups, the percentages of (P), (S), and (PS) MLK were relatively high. This suggests that both the NS and NNS-H students utilized (P), (S), and (PS) MLK substantially when choosing the indefinite article *a/an*. In the use of *Ø*, the results of metalinguistic analyses between the NNS-H and NS groups were quite different. In using *Ø*, the high-proficiency students used (S) MLK at a substantial rate, followed by (PS) MLK. The NS students, on the contrary, used (P) MLK more frequently than (S). The combination (PS) MLK was used in relatively small percentages by the NS group. For the NNS-L students, the results showed that they used (S) MLK at the highest percentage for each article in all NP types.

Another notable point was that non-explicit explanations were found at a much higher rate in the NS group than in the two NNS groups, particularly with the use of Ø in all contexts. Among the Thai students, the low-proficiency group supplied more non-explicit explanations than the students in the high-proficiency group.

According to the quantitative analyses conducted herein, it was found that the findings supported hypotheses 1 and 2. While the MLK used by the native speakers was pragmatically-oriented, the MLK used by the Thai EFL learners was primarily semantically and syntactically-oriented. The results revealed that the MLK used by the high-proficiency students was more pragmatically-oriented than the low-proficiency students, which proved that the MLK used by the three participant groups was different pragmatically, but somehow similar syntactically and semantically.

In testing the third hypotheses, the metalinguistic explanatory task was further analyzed qualitatively to determine the participants' mental representations when they made their article choices on the test. The results revealed that with respect to the generic use of articles, misunderstanding of the generic use of articles was notable. Most of the correct responses by the low-proficiency students were chosen based on inappropriate reasons (i.e. without incorporating generic interpretations), or simply guessing. In other words, the proper article was selected without the students actually understanding the generic meaning of the context. In a generic NP environment, what was normally occurring was that the students frequently looked at the countability and singularity of the NP in question and chose *a/an* simply because the NP was deemed to be singular. This would explain why the seemingly accurate use of *a/an* in the indefinite generic context among the low-proficiency students came to a relatively high percentage (65.0%). The high-proficiency students were more capable of recognizing generic readings than the low-proficiency students. However, they tended to be predisposed toward using *a/an* when a given NP was in the form of a singular count noun. Among the high-proficiency students, the use of '*the* + a singular count noun' was strongly connected with Type 2 [+SR, +HK] article usage rather than Type 1 generic use. The native English speakers, in contrast, used *the* more frequently in generic contexts regardless of the existing subtle difference between generic-*a/an* and generic-*the* uses. As a result, the percentage of using *a/an* in the generic-*a/an* context in the native speaker group decreased to only 54.2%, which was lower than both groups of Thai students.

With regard to performance on other NP types, the results revealed that the students in the low-proficiency group had more difficulty in accurately detecting referentiality, countability and number than the students in the high-proficiency group. The reasons why the students had problems with detecting these notions varied depending largely on their levels of English proficiency and differences in the students' metalinguistic knowledge of the English article system. Typically, for Thai learners, metalinguistic knowledge is developed through the initial input (i.e. instructional rules) provided by teachers and textbooks. It was found that many Thai students in this study alluded to a number of basic rules which they had likely been taught in school. The students, particularly those with low English proficiency, usually adhered to the basic instructional rules in explaining their article choices. Roughly speaking, these rules seem to correspond to the notions of SR, HK, countability and number. However, the problem is that the students often applied these rules separately without understanding the relationships that exist among these important features, which makes the rules inter-dependent on each other. The hypotheses for using *the* and *a/an* were usually formulated as: 'specific = *the*,' or 'a singular count noun = *a/an*'. Thus, some students were confused when a reference was both singular and specific, or when a reference was plural and non-specific.

From the students' explanations, the frequency of failing to consider referentiality was notable among the students in the low-proficiency group. With limited lexical knowledge, the students with low proficiency often lacked the ability to comprehend global contexts. As a result, they had to rely heavily on limited structural cues—such as the existence of the plural morpheme—in order to determine articles. In their attempt to use articles, these students formed various non-generalizable rule-based hypotheses including word-article collocational rules. The most frequently-mentioned collocational rules were those involving prepositions and certain articles. In the high-proficiency group, the students expressed more understanding of the importance of referentiality in properly using articles. Such students also recognized that referentiality was deeply embedded in contexts. As these students were more capable of comprehending contexts, they tried to associate referentiality in terms of larger contextual dynamics. This supports why the high-proficiency students supplied *the* in Type 2 items more properly and with more linguistically-appropriate reasons than the low-proficiency students.

Their explanations were found to be based primarily on pragmatic considerations with regards to referentiality of the given NPs.

However, that the high-proficiency students depended more on referentiality does not mean that they could correctly detect whether or not a referent was assumed to be known or identifiable by the hearer [HK]. What was found to be the case was that the NNS-H students often assumed that whatever was 'specific' to them should be treated as if it was identifiable by the hearer/reader. When the students brought too much presupposition and external knowledge to their readings, such knowledge often interfered with making an accurate judgment about when *the* should be used. Many learners tended to assume that once they could identify the referent (perhaps from their own presupposed beliefs), then *the* was the proper article. In fact, this is not always the case because 'specific reference' and 'assumed hearer's knowledge' are two separate notions. The students who did not clearly understand the notion of HK tended to mark first-mention nouns with *the* based on their assumption that whatever was identifiable by them would also be identifiable by the hearer. The students' misconception about SR and HK frequently resulted in the overgeneralization of *the* with [-HK] nouns.

It should be noted that the term 'specific' is often used in basic instructional rules to refer to the condition that calls for the use of the definite article in both L1 and EFL article lessons. The unclear explanation of the term seems to prevent Thai students from developing an actual understanding of the distinction between SR and HK. The lack of distinction between SR and HK was a major source for misdetection of HK, which is the primary requirement for determining the use of *the*. From the students' explanations, it was noted that the notions of SR and HK were often mixed up. The typical use of the words 'specific' and 'non-specific' among the students was mainly to distinguish between the usage of the definite article and the indefinite article. It appears that in some cases the students used 'specific' to describe 'second- or subsequently-mentioned' nouns, while in other cases they used the term to mean 'identifiable'. However, in such cases it was often ambiguous whether they considered the referent to be 'identifiable' by the speaker [+SR] or the hearer [+HK]. This conceptual misunderstanding that the two notions were undistinguishable seemed to be one of the major reasons for the Thai students' inability to use English articles properly.

In fact, many of the high-proficiency students expressed an understanding that HK is a separate entity from SR. These students, thus, did not limit their attention only to limited local contexts but extended it to more global contextual dynamics. As a result, they faced a different difficulty in using articles. The acknowledgment of the importance of dynamic context-based judgments for article choices did not guarantee the proper detection of HK among the high-proficiency students. This detection turned out to be problematic for them; they found it difficult to judge what circumstances would actually make a reference identifiable for the hearer. The misdetection of HK represented a large portion of the high-proficiency students' overuse of the definite article 'the'. This suggests that pragmatic constraints are more influential than syntactic and semantic constraints in the interlanguage development of the Thai participants. As pragmatic errors are likely the result of the increased information processing burden, they usually do not disappear even at an advanced stage of the L2 acquisition process.

Related to the above problem, there were also cases when the students (even the native English speakers) associated a NP that was semantically-related in some way with an object/person/event that had been mentioned earlier in the text. Consequently, the reference was treated as if it was 'known' to the hearer, and thus *the* was chosen. Many explanations for the choice of *the* showed that the students sometimes failed to detect what should actually be presupposed as assumed hearer's knowledge. Some students used *the* simply because they saw the same word appear earlier in the text, or conversely, they decided not to use *the* if the reference in question did not appear previously in the same text. Other students perceived the existence of modifiers such as preposed adjectives and post-modifying relative clauses as indicators of whether or not the reference in question was identifiable. To many students, the existence of a modifier for a given reference became a convenient symbolic indicator that resulted in automatic insertion of the definite article *the*.

Another notable point is that there was a large decrease in the occurrence of non-generalizable hypotheses and collocations as the students' proficiency level increased. At the low-proficiency level, the students' decisions were based on idiosyncratic or non-generalizable rules and collocations, misunderstanding of syntactic structures and lexicon, and so forth. Most high-proficiency students' decisions for selecting articles were based more on appropriate reasoning than those with low English proficiency. This

suggests that the accurate article choices among the students with high-proficiency contained a greater number of answers based on a more solid understanding of article usage when compared with the students in the low-proficiency group. Moreover, although some non-generalizable rules were formulated among the high-proficiency students, these rules appeared to be commonly-observed across students and were less individually-based than those formulated in the low group.

Interpreting the students' performance in Type 3 [+SR, -HK] and Type 4 [-SR, -HK] contexts marked with *a/an* and *Ø* was more complicated than the other two NP types. This is due to the fact that making proper article choices for items in these two NP types also depends on the proper detection of countability and number of given NPs in addition to referentiality. Accordingly, the students' explanations varied depending on these factors and they did not always rely upon referentiality in their choices of articles. As such, the results of the metalinguistic explanations raise a question as to the validity of comparing the participants' performance across Huebner's NP types which were based solely on referentiality features. As we have seen, the students with low proficiency seldom expressed their reliance upon referentiality, particularly for their choices of *a/an* and *Ø*. Their decisions were mainly based on the detection of countability and number, which indeed are also crucial requirements for the accurate use of both articles. There was evidence that noun countability, which is a means of distinguishing *a/an* from *Ø*, was one problematic factor hindering the proper usage of English articles among a great number of Thai students in this study. Many students often treated countability as if each noun belonged to a static list of countable and uncountable nouns and failed to consider larger dynamic contexts. Thus, countability seems to remain problematic among Thai students, even those with high English proficiency.

Although this study treated *a* and *an* as one article (as they share the same semantic and pragmatic functions), the results revealed that most participants recognized the distinction in the use of the two forms. It was found that the failure to detect a vowel/consonant distinction was relatively rare across groups compared to other features. The students' misuse of *an* (as opposed to *a*) occurred much less frequently than the misuse of other articles. This may be because most students were aware of this phono-syntactic rule as it is often explicitly provided in EFL article instruction. Many students

were influenced by this rule so much that they tended to overlook all other features that should have been detected when making article choices.

Another problem with the instructional rules is that the *zero* article (*Ø*) is not usually described explicitly in EFL lessons, except in the case of idiomatic expressions and conventional uses that require no article. From the students' metalinguistic explanations, it was evident that the students did not have a clear understanding of how the *zero* article (*Ø*) functions. For those items where *Ø* was the correct article, many students in both NNS groups wrongly chose either *a/an* or *the*. Even when the students decided to use *Ø*, they had difficulty explaining why they thought *Ø* was an appropriate article. As indicated by the students' performance, *Ø* constituted the lowest percentage of accurate responses on the test. Moreover, the choice of *Ø* comprised a high proportion of non-explicit reasons (such as guessing). This gives credence to the observation that the use of *Ø* is most problematic among Thai students.

As can be seen, the high- and low-proficiency students had different difficulties in detecting referentiality. The students with high English proficiency did not have a clear understanding of the distinction between the two referentiality features HK and SR, and thus often misdetected HK. The students in the NNS-L group, on the other hand, frequently failed to detect both referentiality features altogether. A large number of students in both groups wrongly considered structural cues such as the existence of modifiers as a device to mark definiteness. The results, thus, supported the third hypothesis that the problems of English article usage among Thai learners are mostly pragmatically-oriented. Moreover, the problems also appeared to arise from the surface syntactic structures of English noun phrases.

7.2 Implications of the Study

7.2.1 Theoretical Implications

As I mentioned in the literature review, native English speakers seem to use the articles easily and smoothly at an early age because they grasp the principles underlying article usage through their L1 acquisition processes. EFL learners, in contrast, have to learn and understand the article system through conceptual analysis. The overall results of the present study demonstrated that pragmatic constraints (i.e. the detection of [±HK])

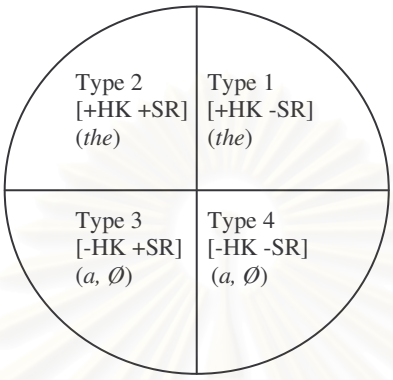
were more influential than semantico-syntactic constraints in the interlanguage development of the Thai participants. The reason is that English articles are not only grammatical elements that follow linguistic rules but are also deeply context-dependent. As the articles maintain contextual links, the discourse aspect—namely HK—appears to be the primary referentiality feature for determining articles.

Although Huebner's (1983) framework labeled the feature SR before HK in the four basic NP types (i.e. Type 1 [-SR, +HK], Type 2 [+SR, +HK], Type 3 [+SR, -HK] and Type 4 [-SR, -HK]), it may be noticed from the organization of these four types that the dominant feature is HK, not SR (i.e. Types 1 and 2 sharing [+HK]; Types 3 and 4 sharing [-HK]). Thus, to show that HK is the prominent feature that underlies these four pragmatic-semantic types, it might be advisable for HK to precede SR—i.e. Type 1 [+HK, -SR], Type 2 [+HK, +SR], Type 3 [-HK, +SR] and Type 4 [-HK, -SR].

Moreover, as suggested by the MLK explanations given by a number of native speakers in the study, there is no clear-cut distinction between generic and non-referential uses of the articles *a/an* and *Ø*. This is in line with Werth's (1980) assertion that generic-*a/an* is not a true generic. The so-called generic-*a/an* should have the same semantic representation as the non-specific. This is also supported by Master's (1990) argument that non-referentials are eventually generic. In fact, as generic-*a/an* and *Ø* denotes a representative example of all entities of a class or species, its reference should be unidentifiable by or unknown to the hearer [-HK]. If this is the case, only generic-*the* should then be classified as Type 1 [+HK, -SR], as its reference as a class or species is often assumed as known to the hearer [+HK]. This will result in generic-*the* and Type 2 [+HK, +SR] article sharing the [+HK] feature. In fact, one may notice the similarity between the two types, as both types exploit the speaker-hearer shared knowledge largely from the context. The difference is in the physical domain and abstract domain of instantiation. The physical domain can be construed as a 'real or actual' object, whereas the abstract domain has to do with the idea or concept of the object as a 'type, class, or species'. When a speaker refers to one particular type or class as a whole, it can be construed as uniquely identified and therefore definite. In this sense, the notion of definiteness can indeed be used for both definite-*the* and generic-*the*. [+HK] is, therefore, distinguishable from [-HK] in terms of definiteness and this feature should distinguish

the article *the* from *a/an* and *Ø*. As such, the model may be modified according to Figure 7.1 below.

Figure 7.1 Pragmatic-Semantic NP Types based on Huebner's (1983) Model



7.2.2 Pedagogical Implications

The findings of this study have led to a few implications for teaching English articles to Thai students. The first pertains to the importance of the initial instruction on English article usage. Once the students have learned (or memorized) a set of basic rules of English article use, it is often difficult (if not impossible) for them to abandon what they have stored in their memory (i.e. as the internalized rule system). Evidence from the present study showed that these initial instructional rules were influential on the students' judgment of article uses, particularly among the low-proficiency students, whose decisions were heavily constrained by rules. Indeed, it is generally assumed that teaching article usage to beginning EFL students is difficult because they still have little lexical knowledge. Students' limited ability to comprehend larger dynamic contexts is likely to hinder the proper detection of referentiality necessary for determining articles. However, with the introduction of lexicon in various contextualized texts appropriate for a certain proficiency level, teachers may systematically incorporate proper article instruction into their EFL lessons. Lack of explicit article instruction at an early stage may greater hinder article usage and is likely to result in fossilization. As students will eventually need to use articles correctly in more advanced studies, it seems impractical to overlook the importance of an initial introduction of how the article system works. While any such instruction might not be very effective at the beginning levels, it could gain greater effectiveness once the students are exposed to more and more English.

Second, the findings suggest that greater attention should be given to the pragmatic domain when teaching English articles. As articles are deeply context-dependent, the use of articles should be introduced through comprehension tasks using discourse grammar. Generally, beginning EFL learners are not fully aware of the differences in grammatical structures between L1 and L2. As a result, they tend to apply the syntactic patterns of their L1 to L2 when they produce the target language. Thus, comprehending a contextualized text may allow the flow of thought to be freer from the interference of L1 than a production task. With increasing proficiency, learners can be more prepared to focus on production activities. When learners begin to produce the target language, constant monitoring is needed.

As MLK should reflect whether students have a true understanding of article usage, the teacher should raise the students' awareness of meta-learning so that they are capable of monitoring their own production tasks through metalinguistic awareness. Swain and Lapkin (1995) suggest that grammatical analysis is essential to accurate production. Learners who lack explicit grammar knowledge will have difficulty understanding the structure of a language.

The third implication is that of the three articles, the definite article *the* tends to be the most unmarked one, for it can be used with almost all nouns in English. The definite article requires the detection of referentiality [±HK] as the major requirement, whereas *a/an* and *Ø* require the proper detection of countability and number in addition to referentiality. Thus, determining the use of indefinite articles *a/an* and *Ø* is likely to be a more complex selective process. The acquisition of grammatical morphemes should proceed from unmarked items to more marked ones. Once learners internalize the use of *the*, they tend to be readier for the distinction between *the* and *a/an*.

The concept of the *zero* article (*Ø*) is more abstract and more difficult to grasp. The use of the *zero* article may be presented last in the sequence.

To sum up, the following points should be considered when teaching English articles to EFL learners:

(1) Contextualized texts should be used to provide language in discourse, as article usage is deeply context-dependent;

(2) A clear conceptual explanation should be provided with regard to the semantic and pragmatic functions of the English articles;

(3) Learners should have a clear understanding of the distinction between the notions of HK and SR. Terminology used to explain the concepts (e.g. 'specific', 'non-specific' and so forth) should be carefully selected to avoid conceptual misunderstanding;

(4) Learners should also have a clear understanding as to how the concepts of HK, SR, countability, and number are inter-dependent of each other. In other words, how article rules are related to each other in determining an appropriate article;

(5) Systematic instruction of the usage of *the*, *a/an*, as well as *Ø* should be provided; and

(6) The concept of noun countability should be introduced in such a way that learners understand how most English nouns can be used as 'count' or 'non-count' depending on the speaker's conceptualization and intended meaning.

Based on the findings of the study, the following is a suggested model to teach the English articles. Four crucial characteristics of the article system are recommended to be introduced to EFL classes:

(1) definite vs. non-definite ([+HK] vs. [-HK])
(2) count vs. non-count ([+count] vs. [-count])
(3) singular vs. plural ([+sing] vs. [-sing])
(4) generic reference

(1) The first step is to decide whether a noun is used in a definite or non-definite sense. The notion of [±HK] should be introduced. If a given referent is [+HK], the noun requires *the*; if [-HK], it requires either *a/an* or *Ø*. To uncover the sense in which a noun is used, teachers may refer to a list of definite descriptions and provide ample examples of contexts for the definite article *the* (e.g. Hawkins's theory: previously-mentioned, immediate situation, larger situation, assumed common knowledge, etc.).

The crucial point is that definiteness is treated before countability. This is because a noun always takes *the* when it is used in a definite sense. This can reduce the student's burden in having to deal with the complex aspect of countability. Moreover, as

Thai has a different way of marking definiteness, it can be assumed that Thai EFL learners are not greatly familiar with this notion. Additionally, *the* is a function word that is used with the highest frequency in the English language. Thus, it seems practical that definiteness is given a foremost attention in article lessons.

(2) When a noun is used in a non-definite sense, the second article characteristic [±count] is applied. The selective process for indefinite articles *a/an* and *Ø* takes more steps than definite article use. The concept of countability itself is highly complicated as most English nouns can be used as count or non-count depending on the speaker's way of expressing objects in terms of discreteness or continuity. Teachers need to discuss the notion of countability and give examples of how most nouns can alternate between mass and count in actual language use.

With the introduction of [±count] at the second stage, a distinction between *a/an* and *Ø* can now be drawn. If a given NP is used as [-count], the NP takes *Ø*; if [+count], it takes either *a/an* or *Ø* depending on the third characteristic, which will be introduced in the next step.

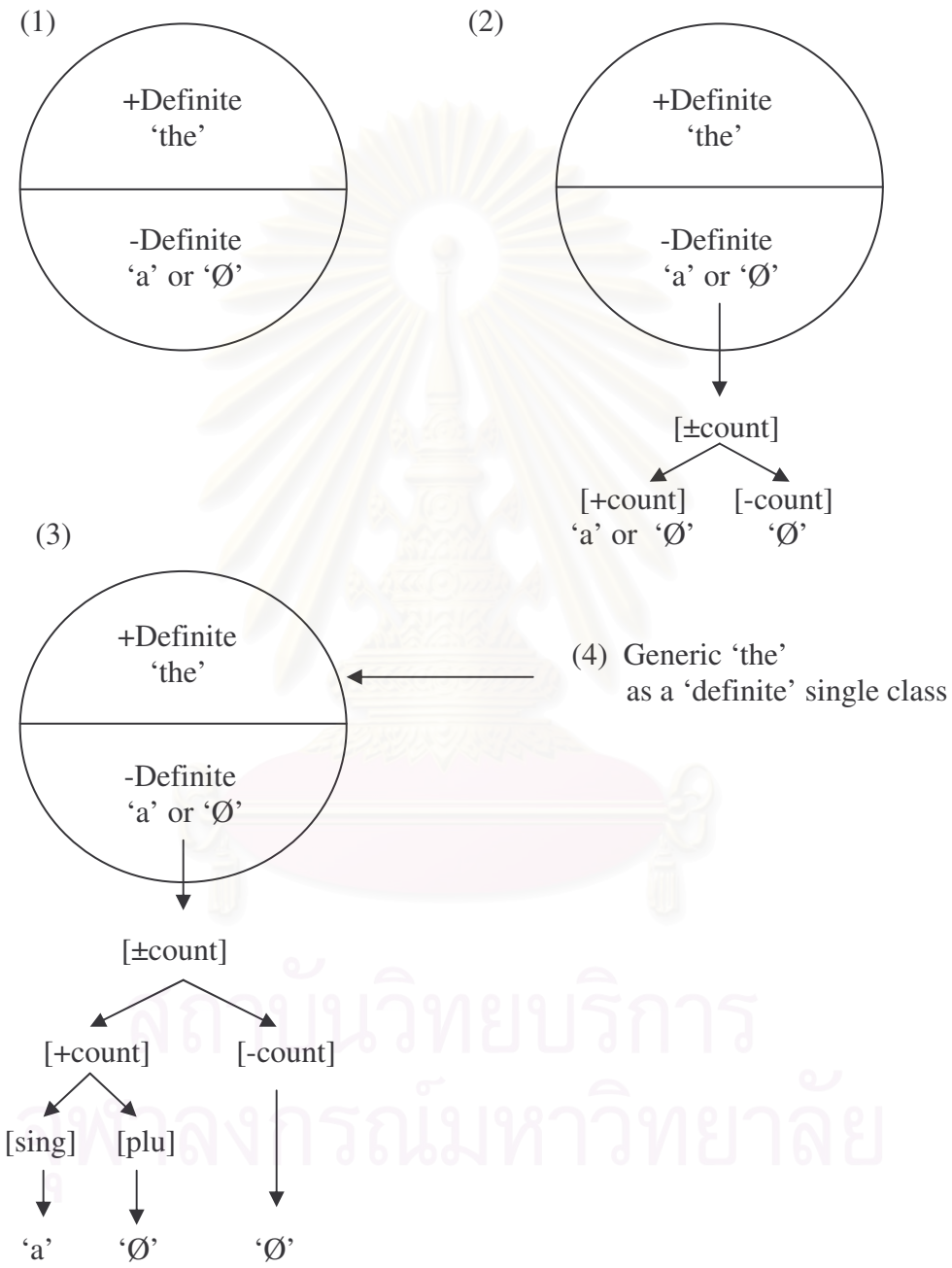
(3) When a noun is [+count], it can be in either a 'singular' or 'plural' form. A singular count noun takes *a/an*; a plural count noun takes *Ø*. At this stage, teachers can introduce the phono-syntactic rule that applies for the distinction between *a* and *an* as well as how nouns are marked in their plural forms.

Toward the end of this stage, teachers may illustrate the subtle differences in meaning between the indefinite descriptions and generic statements using *a/an* and *Ø*. However, as earlier mentioned in the previous chapter, interpretations of indefinite and generic references are difficult to distinguish, for it seems that both are non-definite. This approach, thus, tends to ignore this distinction at least during the beginning and intermediate levels. Such a distinction may be handled later in advanced EFL classes.

(4) As the subtle distinction between indefinite article uses and generic-*a/an* and *Ø* uses tends to be minor enough to ignore, only generic-*the* will be introduced as the fourth article characteristic. As discussed in the previous section, the notion of definiteness can indeed be used for both definite-*the* and generic-*the*, while generic-*a/an* and *Ø* can be classified as non-definite. This is to keep the model as simple as possible.

Figure 7.2 below is a proposed model for introducing English articles in EFL classes using pragmatically-oriented instruction.

Figure 7.2: *A Proposed Model for Teaching English Articles*



Example sentences illustrating the above model:

(1) Definite – *the* (the distinctions of countability and number are neutralized)

Non-count: Pass me the beer.

Countable singular: Have you fed the cat?

Countable plural: The cats have already eaten.

(2) Non-definite – *a/an* and *Ø* (incorporation of Type 3, Type 4, and Type 1)

Non-count: There is beer in the glass.
I drink milk every day.
Electricity is a wonderful helper in the home.

Countable singular: Dad gave me a cat.
You'd better see a doctor.
A fly likes to rub its legs together.

Countable plural: He keeps sending me Ø messages.
Ø Bees carry pollen from Ø flowers.
Ø Mammoths were animals of the elephant family.

(3) Generic – *the* (as a 'definite' single class/group/species)

Countable singular: The penguin lives near the South Pole.

In summary, the model simply introduces the distinction between the discourse feature [±HK], which distinguishes the use of the definite article *the* from the indefinite articles *a/an* and *Ø*. As English marks nouns for definiteness and non-definiteness, not specific reference [±SR], specificity may be dealt with later in advanced EFL lessons. For non-definite references, countability and number are incorporated to distinguish between the two indefinite articles *a/an* and *Ø*. As usage of indefinite articles is eventually generic, the distinction between non-definite and generic contexts is not needed until later in advanced EFL lessons. Only generic-*the* is introduced in terms of a definite class or species of entities as a whole. As learners start to fully understand how the article system works, instruction can be given about the subtle difference between the three generic article uses: *a/an*, *Ø*, and *the*. A comparison can also be drawn between generic interpretations and the general notions of definiteness and non-definiteness that learners have learned in earlier lessons.

7.3 Recommendations for Further Research

This study offers an alternative approach to address the problems in acquiring the English article system among Thai learners in EFL contexts. It investigated Thai learners' metalinguistic knowledge with regard to English article usage in their interlanguage development. The information provided herein reveals the problems which Thai learners of different proficiency levels have in learning the English article system. Suggestions have been made about aspects of English articles and how these aspects can be introduced in EFL classes. Based on the main findings of the study, the following recommendations are made for further research.

Firstly, this study investigated the participants' performance on English article usage, using an objective test of English articles. Therefore, continued research is needed to investigate students' metalinguistic knowledge in the acquisition of English articles, using a production task such as essay writing. A longitudinal study of a small group of Thai learners with different proficiency levels may be performed in order to determine to what extent the students' spontaneous monitoring of their production tasks through metalinguistic awareness (facilitated by the teacher) would lead to improved acquisition and development of the students' performance on article usage.

Secondly, as noted in the findings of the present study, noun countability was another feature that constituted a major obstacle in correctly using articles. Noun countability remained problematic even for high-proficiency Thai students in the study. The difficulty comes from the fact that most English nouns can be used as either count or non-count depending on intended meaning and the surrounding context. Countability judgments should be based on the speaker's conceptualization (as discussed in Chapter 2) rather than arbitrary lists of countable and uncountable nouns. Thus, to use articles properly, countability detection should go beyond the lexical-head level to the whole-NP level. As the framework used in the present study focuses on referentiality, not countability, further research may be undertaken to investigate how countability judgments in English discourse affect English article use. A future study may also explore how article instruction could be made more effective if various exercises in countability detection were incorporated so that the students can experience how native

speakers change their perception of countability depending on the intended meaning and the context.

Finally, instructional materials can be developed based on this simplified framework for teaching English articles and further research can be conducted to provide statistical results of the applications as empirical evidence.

REFERENCES

Abbott, B. (2004). Definiteness and indefiniteness. In L.R. Horn & G. Ward (Eds.). *The Handbook of Pragmatics*. Blackwell Publishing.

Agnihotri, R.K, Khanna, A.L. & Mukherjee, A. (1984). The use of articles in Indian English: Errors and pedagogical implications. *IRAL*, 22: 115-29.

Akiyama, M.M. & William, N. (1996). Spatial components in the use of count nouns among English speakers and Japanese speakers of English as a second language. *Language Learning*, 46/2: 217-31.

Allan, K. (1980). Nouns and countability. *Language*, 56: 541-67.

Arthur, B. (1980). Gauging the boundaries of second language competence. A study of learner judgements. *Language Learning*, 30: 177-194.

Bailey, N., Madden, C. & Krashen, S. (1974). Is there a "natural sequence" in adult second language learning? *Language Learning*, 24/2: 235-43.

Balwin, T. & Bond, F. (2003). *Learning the countability of English nouns from corpus data*. Retrieved from www.lingo.standard.edu/pubs/tbaldwin/acl2003.pdf.

Bataineh, R.F. (2005). Jordanian undergraduates EFL students errors in the use of the indefinite article. *Asian EFL Journal*, Vol. 7/1. Retrieved from http://www.asian-efl-journal.com/march_05_rfb.php.

Berry, R. (1991). Re-articulating the articles. *ELT Journal*, 45/3: 252-59.

Bialystok, E. (1979). Explicit and implicit judgments of L2 grammaticality. *Language Learning*, 29/1:81-103.

Bialystok, E. (1982). On the relationship between knowing and using linguistic forms. *Applied Linguistics*, 3/3:181-206.

Bialystok, E. (1990). The competence of processing: Classifying theories of second language acquisition. *TESOL Quarterly*, 24: 635-648.

Bialystok, E. (1994). Representation and ways of knowing: Three issues in second language acquisition. In N. Ellis (Ed.). *Implicit and explicit learning of languages*. London: Academic Press, pp. 549-569.

Bickerton, D. (1981). *Roots of Languages*. Ann Arbor, Michigan: Karoma Press.

Birdsong, D. (1989). *Metalinguistic Performance and Interlinguistic Competence*. Berlin: Springer-Verlag.

Bley-Vroman, R. (1986). Hypothesis testing in second-language acquisition theory. *Language Learning*, 36/3:353-76.

Bolinger, D. (1977). *Meaning and Form*. London: Longman.

Brown, J.D. (1983). An Exploration of morpheme-group interactions. In K.M. Bailey, M.H. Long, & S. Peck (Eds.), *Second language acquisition studies* (pp. 25-40). Rowley, MA: Newbury House Publishers.

Brown, G. & Yule, G. (1984). *Discourse Analysis*. Cambridge University Press.

Butler, Y. (1999). *The role of metacognition in the development of the English article system among non-native speakers*. Unpublished doctoral dissertation, Stanford University.

Butler, Y. (2000). *Japanese Students' Metalinguistic Knowledge and Strategies in Their Use of English Articles*. Presented in The Second Annual Conference of the Japanese Society for Language Studies (JSLS) in August, 2000.

Butler, Y. (2002). Second language learners' theories on the use of English articles. *Studies in Second Language Acquisition*, 24/3: 451-80.

Carpenter, B. (1997). *Type-Logical Semantics*. Cambridge, Massachusetts: The MIT Press.

Celce-Murcia, M. & Larsen-Freeman, D. (1999). *The Grammar Book: An ESL teacher's Course*. Boston: Heinle & Heinle.

Chafe, W. (1976). Givenness, contrastiveness, subjects, topics, and point of view. In Li, C.N. (Ed.). *Subject and Topic*. New York: Academic Press, pp. 23-55.

Chaney, C. (1989). I pledge a legiance to the flag: Three studies in word segmentation. *Applied Psycholinguistics*, 10: 261-281.

Chaney, C. (1994). Language development, metalinguistic awareness, and emergent literacy skills of 3-year-old children in relation to social class. *Applied Psycholinguistics*, 15/3: 371-94.

Chaudron, C. (1983). Research on metalinguistic judgments: A review of theory, methods, and results. *Language Learning*, 33/3: 343-77.

Christophersen, P. (1939). *The Articles: A Study of Their Theory and Use in English*. Copenhagen: Munkgaard.

Correa, M. (2004). *Literacy, metalinguistic awareness and multi-competence*. University of Arizona.

Cruse, A. (2004). *Meaning in Language: An Introduction to Semantics and Pragmatics*, Second Edition. Oxford: Oxford University Press.

Cziko, G.A. (1986). Testing the language bioprogram hypothesis: A review of children's acquisition of articles. *Language*, 62/4: 878-98.

DeKeyser, R.M. (1995). Learning second language grammar rules. *SSLA*, 17, 379-410.

DeKeyser, R.M. (1997). Beyond explicit rule learning: Automizing second language morphosyntax. *SSLA*, 19: 195-221.

DeKeyser, R.M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, 55/1:1-25.

Dulay, H. & Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning*, 24/1: 37-53.

Dulay, H., Burt, M. & Krashen, S. (1982). *Language Two*. Oxford: Oxford University Press.

Ekiert, M. (2004). Acquisition of the English article system by speakers of Polish in ESL and EFL settings. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*, Vol. 4, No. 1. Retrieved from http://journals.tc-library.org/ index.php/ tesol/article/ view PDF Interstitial/42/49.

Ellis, N.C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *SSLA*, 24: 143-188.

Ellis, N.C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *SSLA*, 27, 305-352.

Ellis, R. (1985). Sources of variability in interlanguage. *Applied Linguistics*, 6:118-31.

Ellis, R. (2002). Does form-focused instruction affect the acquisition of implicit knowledge? *SSLA*, 24, 223-236.

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, 54/2:227-275.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language. *SSLA*, 27: 141-172.

Feigenbaum, I. (1985). *The Grammar Handbook*. New York: Oxford University Press.

Galambos, S.J. & Hakuta, K. (1988). Subject specific and task-specific characteristics of metalinguistic awareness in bilingual children. *Applied Psycholinguistics*, 9: 141-162.

Gass, S. (1983). The development of L2 intuitions. *TESOL Quarterly*, 17/2:273-291.

Goldschneider, J.M. & DeKeyser, R.M. (2001). Explaining the "Natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning*, 51/1: 1-50.

Gombert, J.E. (1992). *Metalinguistic development*. Chicago: The University of Chicago Press.

Grannis, O.C. (1972). The definite article conspiracy in English. *Language Learning*, 22: 275-89.

Green, P. & Hecht, K. (1992). Implicit and explicit grammar: An empirical study. *Applied Linguistics*, 13: 168-184.

Hakuta, K. (1976). A case study of a Japanese child learning English as a second language. *Language Learning*, 26/2: 321-51.

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: A Study in Reference and Grammaticality Prediction*. London: Croom Helm.

Hawkins, J.A. (1980). On surface definite articles in English. In Johan Van der Auwera (Ed.). *The Semantics of Determiners*. Baltimore: University Park Press.

Hawkins, J.A. (1991). On (in)definite articles: Implications and (un)grammaticality prediction. *Journal of Linguistics*, 27: 405-442.

Hornby, A.S. (1974). *Oxford Advanced Learner's Dictionary of Current English*. Third Edition. London: Oxford University Press.

Horowitz, F.E. (1989). ESL and prototype theory: Zero vs. definite article with place names. *IRAL*, 27: 81-98.

Hu, G. (2002). Psychological constraints on the utility of metalinguistic knowledge in second language production. *SSLA*, 24: 347-386.

Huebner, T. (1979). Order-of-acquisition vs. dynamic paradigm: A comparison of method in interlanguage research. *TESOL Quarterly*, 13: 21-28.

Huebner, T. (1983). *A longitudinal analysis of the acquisition of English*. Ann Arbor, MI: Karoma.

Huebner, T. (1985). System and variability in interlanguage syntax. *Language Learning*, 35/2: 141-63.

Hulstijn, J.H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning. *SSLA*, 27, 129-140.

Hulstijn, J.H., & Hulstijn, W. (1984). Grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning*, 34: 23-43.

Iamchote, D. (1971). *The definite and indefinite articles in English and their Thai equivalents*. M.A. Thesis, Chulalongkorn University.

Ionin, T. & Wexler, K. (2002). The certain uses of *the* in L2 English. In J.M. Liceras, H. Zobl & H. Goodluck (Eds.). *Proceedings of the 6th Generative Approaches to Second Language Acquisition Conference (GASLA 2002): L2 Links*. Cascadilla Press. Retrieved from http://www.lingref.com/cpp/gasla /6/paper 1039.pdf.

Isarankura, S. (2007). The acquisition of English articles by Thai learners. In *Proceedings of the 7th International Symposium on Natural Language Processing (SNLP 2007)*.

Isarankura, S. (2008). Metalinguistic knowledge in English article acquisition among Thai learners. In *Proceedings of the 13th Conference of Pan-Pacific Association 'of Applied Linguistics (PAAL 2008)*.

Jarvis, S. (2002). Topic continuity in L2 English article use. *Studies in Second Language Acquisition*, 24/3: 387-418.

Kaku, K. (2006). Second language learners' use of English articles: A case study of native speakers of Japanese. *CLO/OPL Ottawa*, Vol. 34: 63-74. Retrieved from http://aix1.uottawa.ca/~clo/Kaku.pdf.

Kaluza, H. (1963). Teaching the English articles to speakers of Slavic. *Language Learning*, 13/2: 113-24.

Kambou, M.K. (1999). Huebner's (1983) semantic wheel for NP reference and L3 acquisition. *Issues in Applied Linguistics*, 10/1: 57-73.

Kearns, K. (2000). Definite Descriptions. In *Modern Linguistics Series: Semantics*. London: Macmillan Press Ltd.

Kharma, N. (1981). Analysis of the errors committed by Arab university students in the use of the English definite/indefinite articles. *IRAL*, 19/4: 333-45.

Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.

Larsen-Freeman, D. (1975). The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly*, 9/4: 409-30.

Lee, D. (2002). Count and mass nouns. In *Cognitive Linguistics: An Introduction*. Oxford University Press.

Lee, E.H. (2007). Teaching English articles by learners' proficiency levels. *English Language & Literature Teaching*, 13/4: 109-126.

Leech, G. & Svartvik, J. (1994). *A Communicative Grammar of English*, Second Edition. London: Longman.

Leow, R.P. (2000). A study of the role of awareness in foreign language transfer, *SSLA*, 22, 557-584.

Littlewood, W.T. (1981). Language variation and second language acquisition theory. *Applied Linguistics*, 11/2: 150-58.

Liu, D. & Gleason, J.L. (2002). Acquisition of the article *the* by nonnative speakers of English. *Studies in Second Language Acquisition*, 24/1: 1-26.

Lu, F.C. (2001). The acquisition of English articles by Chinese learners. *Second Language Studies*, 20/1:43-78.

Lyons, C. (1980). The meaning of the English definite article. In J. Van der Auwera (Ed.). *The Semantics of Determiners*. Baltimore: University Park Press.

Malakoff, M. (1992). Translation skill and metalinguistic awareness in bilinguals. In R. Harris (Ed.). *Cognitive processing in bilinguals*. Amsterdam: North Holland, pp. 515-530.

Master, P. (1988a). Teaching the English article system, Part I. *English Teaching Forum*, 26/2: 2-7.

Master, P. (1988b). Teaching the English article system, Part II. *English Teaching Forum*, 26/3: 18-25.

Master, P. (1990). Teaching the English articles as a binary system. *TESOL Quarterly*, 24: 461-78.

Master, P. (1997). The English article system: Acquisition, function, and pedagogy. *System*, 25/2: 215-32.

Master, P. (2002). Information structure and English article pedagogy. *System*, 30/3: 331-48.

McEldowney, P.L. (1977). A teaching grammar of the English article system. *IRAL*, 5/2: 95-112.

Mizuno, M. (1999). Interlanguage analysis of the English article system: Some cognitive constraints facing the Japanese adult learners. *IRAL*, 37/2:127-52.

Morris, L. (2003). Linguistic knowledge, metalinguistic knowledge and academic success in a language teacher education programme. *Language Awareness*, 12/2: 109-23.

Mufwene, S. (1984). The count/mass distinction and the English lexicon. In D. Testen, V. Mishra, & J. Drogo (Eds.). *Lexical Semantics*. Chicago: Chicago Linguistic Society.

Myhill, D. (2000). Misconceptions and difficulties in the acquisition of metalinguistic knowledge. *Language and Education*, 14/3:151-163.

Neale, S. (1990). *Descriptions*. Cambridge: The MIT Press.

Odlin, T. (1986). On the nature and use of explicit knowledge. *IRAL*, 24/2:123-44.

Oller, J.W. Jr. & Redding, E.Z. (1971). Article usage and other language skills. *Language Learning*, 21/1: 85-95.

Parrish, B. (1987). A new look at methodologies in the study of article acquisition for learners of ESL. *Language Learning*, 37: 361-83.

Platteau, F. (1980). Definite and indefinite generics. In J. Van der Auwera (Ed.) *The Semantics of Determiners*. Baltimore: University Park Press.

Poesio, M. & Vieira, R. (1998). A corpus-based investigation of definite descriptive use. *Computational Linguistics*, 24/2: 184-216.

Prasithrathsint, A., Hoonchamlong, Y., & Savetamalya, S. (2001). *Grammatical Theories*. Bangkok: Chulalongkorn University Press.

Prince, E.F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.). *Radical Pragmatics*. New York: Academic Press, pp.223-56.

Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1972). *A Grammar of Contemporary English*. London: Longman.

Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1979). *A Grammar of Contemporary English*. London: Longman.

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

Rastall, P. (1995). Definite article or no definite article? *English Today*, 11/2:37-39.

Renou, J.M. (1998). *Effect of grammaticality judgements on the relationship between metalinguistic awareness and second language proficiency*. Ph.D. Dissertation. University of Ottawa, Ontario, Canada.

Renou, J.M. (2001). An examination of the relationship between metalinguistic awareness and second-language proficiency of adult learners of French. *Language Awareness*, 10/4:248-267.

Ricciardelli, L.A. (1993). Two components of metalinguistic awareness: Control of linguistic processing and analysis of linguistic knowledge. *Applied Psycholinguistics*, 14/3: 349-67.

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. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *SSLA*, 18: 27-67.

Roehr, K. (2005). Theorizing L2 metalinguistic knowledge. *Essex Research Report in Linguistics*, Vol. 50:21-70.

Roehr, K. (2006). Metalinguistic knowledge and language-analytical ability in university-level L2 learners. *Essex Research Report in Linguistics*, Vol. 51:41-71.

Rosa, E. & O'Neill, M.D. (1999). Explicitness, intake, and the issue of awareness. *SSLA*, 21: 511-556.

Russell, B. (1905). On denoting. *Mind*, 14:479-93. Reprinted in R.C. Marsh (Ed.). *Logic and Knowledge*. London: George Allen and Unwin.

Searle, J. (1969). *Speech Acts*. London: Cambridge University Press.

Seliger, H. (1979). On the nature and function of language rules in language teaching. *TESOL Quarterly*, 13: 359-369.

Selinker, L. (1972). Interlanguage. *IRAL*, 10/3:209-231.

Sharma, D. (2005). Language transfer and discourse universals in Indian English article use. *Studies in Second Language Acquisition*, 27: 535-566.

Sharwood-Smith, M. (1981). Consciousness-raising and the second language learner. *Applied Linguistics*, 2: 159-168.

Sinclair, J.M. (Ed.), 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Snape, N. (2005). The certain uses of articles in L2-English by Japanese and Spanish speakers. *Durham and Newcastle Working Papers in Linguistics*, 11: 155-68.

Snow, C.E. (1990). The development of definition skill. *Journal of Child Language*, 17/3: 697-710.

Sorace, A. (1985). Metalinguistic knowledge and language use in acquisition-poor environments. *Applied Linguisitcs*, 6/3:239-54.

Srioutai, J. (2001). Viewing and Teaching the English Articles Semantically. *Journal of Arts*, 30/2: 151-167.

Standwell, G.J. (1997). The English articles: A world wide exception? *IRAL*, 35/4: 269-76.

Supol, D. (1980). Articles in Thai learners. *Journal of Humanities*, 10/3:40-58.

Swain, M. & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16/3:371-391.

Tarone, E.E. (1985). Variability in interlanguage use: A study of style-shifting in morphology and syntax. *Language Learning*, 35/3: 373-403.

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: 335-55.

Thonglor, K. (1997). *The Principles of Thai Language*. Bangkok: Ruamsarn.

Tokowicz, N. & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar. *SSLA*, 27: 173-204.

Trask, R.L. (1999). *Key Concepts in Language and Linguistics*. London: Routledge.

Uppakitsilapasarn, Phraya. (2002). *The Principles of Thai Language*, Eleventh Edition. Bangkok: Thai Wattanapanich.

Ward, G. & Birner, B. (1995). Definiteness and the English existential. *Language*, 74/4: 722-42.

Wenden, A.L. (1998). Metacognitive knowledge and language learning. *Applied Linguistics*, 19/4: 515-37.

Werth, P. (1980). Articles of Association: Determiners and Context. In J. Van der Auwera (Ed.): *The Semantics of Determiners*. Baltimore: University Park Press.

Whitman, R.L. (1974). Teaching the article in English. *TESOL Quarterly*, 8: 253-62.

Wierzbicka, A. (1988). *The Semantics of Grammar*. Amsterdam: John Benjamins.

Yamada, J. & Matsuura, N. (1982). The use of the English article among Japanese students. *RELC Journal*, 13/1: 50-63.

Yelland, G.W., Pollard, J. & Mercuri, A. (1993). The metalinguistic benefits of limited contact with a second language. *Applied Psycholinguistics*, 14/4: 423-44.

Yoon, K.K. (1993). Challenging prototype descriptions: Perception of noun countability and indefinite vs. zero article use. *IRAL*, 31: 269-89.

Young, R. (1996). Form-function relations in articles in English interlanguage. In R. Bayley & D.R. Preston (Eds.): *Second Language Acquisition and Linguistic Variation*. Amsterdam: John Benjamins, 135-75.

Yule, G. (1996). *Pragmatics*. Oxford: Oxford University Press.

APPENDICES

APPENDIX A

Test of English Articles

Fill in the blank with ***a/an, the*** or ***Ø***.

I. Of all the animals on earth, (1) Ø man is the most intelligent. Other animals are also able to think, and perhaps (2) the best thinker among them is (3) the dolphin. (4) The dolphin has been (5) a friend of (6) Ø sailors for centuries. It swam alongside (7) Ø ships in the old days and still likes to today. Perhaps it must swim fast, but it still enjoys (8) the thrill of keeping up with (9) a vessel.

(10) Ø scientists have found that (11) Ø dolphins talk to each other, but so far nobody has been able to say exactly what their squeaks and whistles mean. It has only been in the last few years that (12) Ø people have learned to do something that (13) Ø dolphins have always been able to do. By sending out (14) Ø sounds which bounce back like (15) Ø echoes, (16) a dolphin is able to tell whether (17) a rock or fish lies ahead of it. Today we have (18) Ø echo-sounders that work for us in much (19) the same way.

[Adapted from Howard, P. (1990). *Basic Skills: Reading/Comprehension Level 2*. Rose Bay N.S.W.: A Jim Coroneos Publication.]

II. Yesterday evening (20) a fire badly damaged the Grand Hotel in San Bernardino. (21) Ø firefighters worked very hard to stop (22) the fire. The police believed it was started deliberately. They found (23) an empty gasoline can in one of (24) the hotel elevators. (25) Ø broken glass injured a number of the guests who were enjoying (26) a party at (27) the hotel. (28) A German woman cut her arm badly. There was (29) Ø blood everywhere, and she had to be rushed to the hospital in (30) a police car. Police interviewed (31) the guests and hotel staff to find out what had happened. (32) A man told our reporter that he saw (33) Ø suspicious-looking men enter the elevator. He gave (34) Ø detailed descriptions to the police. So far, (35) the hotel manager has refused to comment on (36) the situation.

[Adapted from Evans, V. (1999). *Round-Up: English Grammar Practice 5*. Essex: Pearson Education Limited.]

III. (37) The lion is one of several of (38) the large cats that are graceful but powerful killers. (39) Ø lions gather in groups. Unlike most cat species, (40) a male lion is different from (41) a lioness. His head is much larger than that of (42) a female lion.

When (43) a lion cub is born, it feeds on its mother's milk for (44) the first three months. Then it begins to eat meat brought to it by its parent. By eighteen months it is able to hunt for itself. (45) The lioness does most of the hunting. She stalks (46) an animal by crawling on her belly. Then she makes (47) a sudden short rush, leaping on (48) the prey to bring it down with her paw.

[Adapted from Howard, P. (1990). *Basic Skills: Reading/Comprehension Level 2*. Rose Bay N.S.W.: A Jim Coroneos Publication.]

APPENDIX B - 1

Metalinguistic Knowledge Questionnaire on English Article Use

In the blank provided, please state the rule or your reason for each of the article you have used.

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____
9. _____
10. _____
11. _____
12. _____
13. _____
14. _____
15. _____
16. _____
17. _____
18. _____
19. _____
20. _____
21. _____
22. _____

23. _____

24. _____

25. _____

26. _____

27. _____

28. _____

29. _____

30. _____

31. _____

32. _____

33. _____

34. _____

35. _____

36. _____

37. _____

38. _____

39. _____

40. _____

41. _____

42. _____

43. _____

44. _____

45. _____

46. _____

47. _____

48. _____

APPENDIX B - 2

<u>แบบสำรวจความรู้ด้านอภิภาษาศาสตร์ของนักเรียนไทยในการใช้คำนำหน้านามในภาษาอังกฤษ</u>

กรุณาเขียนเหตุผลหรือหลักไวยากรณ์ที่ท่านใช้ในการตัดสินใจเลือก article ตามคำตอบที่ท่านให้ไว้ในข้อสอบ

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____
9. _____
10. _____
11. _____
12. _____
13. _____
14. _____
15. _____
16. _____
17. _____
18. _____
19. _____
20. _____
21. _____
22. _____

23. _____

24. _____

25. _____

26. _____

27. _____

28. _____

29. _____

30. _____

31. _____

32. _____

33. _____

34. _____

35. _____

36. _____

37. _____

38. _____

39. _____

40. _____

41. _____

42. _____

43. _____

44. _____

45. _____

46. _____

47. _____

48. _____

APPENDIX C - 1

Date_____

Background Questionnaire:

1. Name_____
2. Age_____
3. Gender_____
4. Native country_____
5. Native language_____
6. Language(s) you speak at home_____
7. Education_____
8. Field of study_____
9. Telephone Number_____
10. Email address_____

Thank you very much for your support.

APPENDIX C - 2

แบบสำรวจประสบการณ์การเรียนภาษาอังกฤษ

กรุณาตอบคำถามในแบบสำรวจนี้ให้สมบูรณ์ถูกต้อง **ตรงตามความเป็นจริงมากที่สุด**

ขอขอบคุณอย่างสูงที่ให้ความร่วมมือในงานวิจัยครั้งนี้

ส่วนที่ 1 ข้อมูลส่วนตัว

ชื่อ นาย/นางสาว.....นามสกุล.....รหัสนักศึกษา.....

อายุ.....ปี หมายเลขโทรศัพท์.....Email.....

วันและเวลาที่ท่านสะดวกหากต้องมีการนัดหมายครั้งต่อไป.....

ส่วนที่ 2 ประวัติการเรียนและการใช้ภาษาอังกฤษโดยทั่วไป

1. ท่านเริ่มเรียนภาษาอังกฤษตั้งแต่อายุ.....ปี ขณะเรียนอยู่ชั้น.....
2. ขณะเป็นนักเรียน ท่านเคยเรียนกับอาจารย์ชาวต่างชาติที่เป็นเจ้าของภาษาหรือไม่..... (ถ้าเคย) เป็นระยะเวลารวม .....ปี.....เดือน.....สัปดาห์.....วัน
3. ท่านเคยไปศึกษา / อบรม ในต่างประเทศที่สื่อสารกันด้วยภาษาอังกฤษหรือไม่..... (ถ้าเคย) โปรดระบุวิชาที่ไปเรียน ..... ประเทศ..... เป็นเวลา.....สัปดาห์.....เดือน.....วัน
4. ท่านเคยไปทัศนาจรหรือพักอาศัยในต่างประเทศที่สื่อสารกันด้วยภาษาอังกฤษหรือไม่..... (ถ้าเคย) เป็นจำนวน.....ครั้ง
   - ครั้งที่ 1: ประเทศ..... เป็นเวลา.....เดือน.....สัปดาห์.....วัน
   - ครั้งที่ 2: ประเทศ..... เป็นเวลา.....เดือน.....สัปดาห์.....วัน
   - ครั้งที่ 3: ประเทศ..... เป็นเวลา.....เดือน.....สัปดาห์.....วัน
   - เกินกว่า 3 ครั้ง (โปรดระบุ).....
     .....
5. ท่านเรียนเสริมพิเศษวิชาภาษาอังกฤษหรือไม่..... (หากเรียน) เฉลี่ยสัปดาห์ละ.....ชั่วโมง รวมเป็นเวลา.....ปี.....เดือน
6. ท่านทำงานพิเศษหรือมีกิจกรรมที่ต้องใช้ภาษาอังกฤษบ่อยครั้ง เช่น โต้วาที เล่นละคร สอนพิเศษ เป็นไกด์ พนักงานต้อนรับ แปล-เอกสาร ฯ หรือไม่ โปรดระบุ.....
7. กิจกรรมใด (ทั้งในและนอกชั้นเรียน) ที่ช่วยให้ท่านเข้าใจภาษาอังกฤษหรือใช้ภาษาอังกฤษได้ดีขึ้น..... .....

ส่วนที่ 3 ประวัติการเรียน English articles: a, an, the

1. ท่านเคยเรียน articles กับอาจารย์ชาวต่างชาติที่เป็นเจ้าของภาษาหรือไม่.....
2. หลักในการใช้ articles: a, an, the รวมทั้งการไม่ต้องใส่ article ที่ท่านทราบและจำได้ดี คือ:-
   - ก. .....
   - ข. .....
   - ค. .....
   - ง. .....
   - จ. .....
   - ฉ. .....
3. กฎที่ท่านทราบช่วยให้ท่านเลือกใช้ article ได้ถูกต้องเสมอหรือไม่.....เพราะ..... .....
4. กิจกรรมใด (ทั้งในและนอกชั้นเรียน) ที่ช่วยให้ท่านเข้าใจ articles และใช้ได้ถูกต้องมากขึ้น..... .....

APPENDIX C- 2 (TRANSLATION)

Students' Background Questionnaire (For Thai Students: Translated from Thai)

Part 1: Personal Information

First Name: ……Last Name: ……..……Reg. No.: …….

Age: …….Tel.: ...……Email: ……..

Convenient day and time for future appointment: ……...

Part 2: English Language Learning Experience

1. When did you start learning English? From Grade…….. Age….……

2. Have you studied English with a native English-speaking teacher? Yes/No……

   If yes, for how long? ……

3. Have you been in any English speaking country for studying/training? Yes/No…….

   If yes, what country/countries……...

   How long? ……

   What subject(s) did you study? ……

4. Have you visited or stayed in any English speaking country as a tourist? Yes/No……..

   If yes, please specify below:

   (1) Country……Length of stay…….

   (2) Country……Length of stay…….

   (3) Country……Length of stay…….

   (4) Country……Length of stay…….

5. Besides school instruction, have you had any special English lessons? Yes/No……..

   If yes, how many hours a week? …….For how long? …….

6. Have you engaged in any part-time jobs or activities that require you to use English? (For example, an English speech contest, a debate, a drama, a part-time English teaching job, a part-time tourist guide, a receptionist, a job as a translator, etc.)

   Yes/No…….If yes, please specify……

   ……...

7. What activities (in class or outside the classroom) help you to understand English and use English better? ……...

   ……..

Part 3: Your Experience in Learning the English Articles

1. Have you studied the English articles with a native English-speaking teacher? Yes/No…….
2. What are some basic rules about English article usage that you can remember?
   (1)……..
   (2)……..
   (3)……..
   (4)……..
   (5)……..
   (6)……..
3. Do you think you can always use the article rules that you know every time you make article choices? Yes/No……..Please explain why? …….
   ……...
4. What activities (in class or outside the classroom) help you to understand how to use the English articles better? ……...
   ……...

BIOGRAPHY

Miss Soisithorn Isarankura graduated with a Bachelor of Arts (first-class honors) degree majoring in English from the Faculty of Liberal Arts, Thammasat University, in Bangkok, Thailand. She received a grant from the British Council under the Colombo Plan to study for her master's degree at the University of Leeds, UK and she graduated with a Master of Arts degree from the Department of Linguistics and Phonetics.

Her interests in variability in the students' interlanguage development began when she was teaching mixed ability students in the English Major program of Dhurakij Pundit University. Thus, in 2005, she commenced her doctoral study in the English as an International Language program of the Graduate School at Chulalongkorn University. Her research interests include the acquisition of English articles in the students' interlanguage. In December 2007, she presented her research paper at the 7th International Symposium on Natural Language Processing (SNLP 2007) in Pattaya, Thailand. Her paper was published in Proceedings of SNLP 2007. Later in August 2008, she also presented her study at the 13th International Conference of Pan-Pacific Association of Applied Linguistics (PAAL 2008) held at the University of Hawaii, Manoa, in Honolulu, USA. Her paper appears in Proceedings of PAAL 2008.