EFL LEARNERS' PERCEPTIONS OF GRAMMATICAL DIFFICULTY IN RELATION TO SECOND LANGUAGE PROFICIENCY, PERFORMANCE, AND KNOWLEDGE

by

Li Ju Shiu

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Curriculum, Teaching and Learning
Ontario Institute for Studies in Education
University of Toronto

EFL LEARNERS' PERCEPTIONS OF GRAMMATICAL DIFFICULTY IN RELATION TO SECOND LANGUAGE PROFICIENCY, PERFORMANCE, AND KNOWLEDGE

Doctor of Philosophy, 2011
Li Ju Shiu
Graduate Department of Curriculum, Teaching and Learning
University of Toronto

Abstract

This study investigated grammatical difficulty from the perspective of second language (L2) learners in relation to their overall L2 proficiency and L2 performance and knowledge. The design included the administration of a student questionnaire, an interview, a proficiency test, and oral production and metalinguistic tasks. The proficiency test and questionnaire were administered to 277 university-level Chinese EFL learners in Taiwan. The questionnaire explored learners' perceptions of grammatical difficulty of 20 English grammar features. Thirty of the students who completed the questionnaire met with the researcher individually to complete a grammatical difficulty ranking activity, 2 grammar exercises and 2 stimulated recalls, all of which aimed to further explore why the learners considered the selected features to be more (or less) difficult for them to learn. The oral production tasks were administered to 27 of the students who completed the questionnaire. The metalinguistic task was administered to 185 of the students who participated in the questionnaire survey.

The questionnaire results indicate that, overall, the participants did not perceive the 20 target features to be difficult to learn. Notwithstanding, the ranking results of the questionnaire suggest that learners' perceptions of grammatical difficulty are based on whether the rules to describe the formation of language features are easy or difficult to articulate. The qualitative results show that the learners' perceptions of grammatical difficulty were influenced by several factors including

their L2 knowledge, L2 grammar learning experience, and L1 knowledge, all of which were examined with reference to syntactic, semantic, and/or pragmatic levels. In terms of the relationship between learners' perceptions of grammatical difficulty and their overall L2 proficiency, results show that learners' perceptions of grammatical difficulty do not vary at the syntactic level, but that there is some variation at the pragmatic level. Regarding the relationship between learners' perceptions of grammatical difficulty and their L2 knowledge, results suggest that learners' perceptions of grammatical difficulty vary according to their implicit/explicit knowledge of the features in question; at the explicit knowledge level, the feature perceived to be less difficult to learn is used more accurately, while at the implicit knowledge level, this is not the case.

Acknowledgements

I would not have completed my Ph.D. study journey without the support and help from the people who have been so important to me these past few years. In particular, this thesis would not have been possible without my thesis committee members' guidance and encouragement. Each committee member deserves my heartfelt thanks. First and foremost, I offer my deepest gratitude to Dr. Nina Spada, my supervisor and mentor, who has made her support available in a number of ways. I am grateful for her faith in me and my potential, her intellectual advice on details of my thesis, and her patience in correcting my grammatical mistakes. I appreciate her heartwarming encouragement, which enabled me to tackle difficulties taking place in different stages of my study journey. I admire her great knowledge and enthusiasm for research in second language learning. Because of her, I was able to complete this long yet fulfilling journey. One simply could not wish for a better supervisor.

My sincere appreciation also goes to Dr. Alister Cumming, whose wisdom and expertise contributed to improve the quality of this thesis. I thank Dr. Cumming for his prompt, insightful comments and suggestions, and I just can't help loving his sense of humor. I would also like to thank Dr. Antoinette Gagne for her non-stopping encouragement and useful suggestions. Her big smiles always made me feel welcomed. I am fortunate to have had Dr. Shawn Loewen as my external examiner and Dr. Eunice Jang as my internal examiner. I am heartily thankful for their genuine and generous comments.

I am indebted to many other professors and professionals who contributed to my academic growth during my study, in particular, many thanks to Dr. Merrill Swain, Dr. Sharon Lapkin, Dr. Jim Cummins, and Dr. Olesya Falenchuk. My heartfelt thanks also go to my friends and colleagues at OISE/University of Toronto. Thank you – Lorena Jessop, Wataru Suzuki, Jennifer Wilson, Yasuyo Tomita, Sebnem Yalcin, Antonella Valeo, Paul Quinn, Daphne Lin, Reed Thomas, Christian Chun, Kyoko Motobayashi, and Joy Zhao. Without your friendship and support, I would not have been able to make this far in my study journey.

I would also like to express my thanks to my colleagues and friends in Taiwan. Thanks to my colleagues' support, I was able to take four-year study leave from my teaching position. Special thanks go to my friends Rachel Hong, Fiona Liu, and Yi-Ching Lee. Thank you for your

patiently listening to my difficulties in numerous nights. My deep gratitude also goes to the students and teachers who participated in the study.

I dedicate this work to my loves—my husband Chien-Chung Lin, my daughter Kuan-Chun Lin, and my son Chen-Yuan Lin. It is difficult for me to find words to appropriately express my appreciation and gratitude to my loves. I thank my husband for his unconditional love and full support. Without him, I would not have been able to make this study journey. Thank you my love for taking all the parenting responsibilities during my four-year absence. Thank you for always being there for me in every step of the way. I thank my daughter and my son for being such sweet kids. Thank you my little babies for trying to understand what mommy was passionate for. I appreciate that you never complained about mommy's absence.

Dedication

To my loves -- Chien-Chung Lin, Kuan-Chun Lin & Chen-Yuan Lin

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter One
Introduction

As a university-level English-as-a-foreign-language (EFL) teacher and second language (L2) researcher, I have long been concerned with how to promote the learning of English grammar by university students. Once, in an English listening and speaking class, I heard one student say, "I think that subject verb agreement is easy to learn because we has studied it for a long time." The comment was made by a first-year university student, who had apparently received many years of grammar-based EFL instruction in high school in Taiwan. The comment made me wonder why a feature (subject verb agreement in this case) that was considered easy to learn was still used incorrectly. Several questions followed, including:

- Does the grammar-based instruction students receive in high school contribute to their English learning in university?
- Which grammar features do students perceive as easy, and which as difficult to learn?
- Why do students perceive some grammar features to be easier to learn than others?
- Do students' perceptions of ease or difficulty regarding grammar features vary with their overall English proficiency?
- How do students' perceptions of grammatical difficulty relate to their actual use?

These questions motivate the current study, which is intended to explore *grammatical difficulty* from the perspective of L2 learners.

Before describing the study, the question of how grammatical difficulty can be defined needs to be discussed as it is central to the issues under investigation.

Defining Grammatical Difficulty in Second Language Learning

The second language acquisition (SLA) literature reveals various approaches to defining "grammatical difficulty." Krashen (1982) puts forward an intuitively appealing idea of "easy rule" and "hard rule," but fails to make the distinction explicit. Green and Hetch (1992) distinguish easy rules from hard rules by the extent to which the rules can be articulated.

According to this view, easy rules are those that are consistently and correctly verbalized. The easy rules that Green and Hetch identify in their study include, "those that (1) [refer] to easily recognized categories; (2) [can] be applied mechanically; (3) [are] not dependent on large contexts: for example the morphological dichotomies like *a/an*, *who/which*, straightforward cases of *some/any*, and simple word order" (p. 179). "Hard rules," on the other hand, refer to those that are difficult to identify or verbalize. One example given by Green and Hetch is aspect, the learning of which requires more than the acquisition of simple exhaustive descriptions.

DeKeyser and Sokalski (1996) consider grammatical difficulty in relation to *comprehension* and *production*. They argue that some grammar structures are easy to comprehend, but difficult to produce, whereas others are easy to produce, but difficult to comprehend. For English speakers learning Spanish, conditional forms of a verb exemplify the former case, while Spanish direct object clitics exemplify the latter. Berent (1985) also considers grammatical difficulty in relation to production and comprehension. In a study that compared ESL learners' production and comprehension of different types of conditional sentences, Berent found that past unreal conditionals were the most difficult to produce, but the easiest to comprehend. In contrast, real conditionals were the easiest to produce, but the most difficult to comprehend.

Some researchers (e.g., Hulstijn, 1995) consider grammatical "difficulty" in terms of the inherent "complexity" of the feature and use these two terms interchangeably. Regardless of how "complexity" is defined, it stands to reason that the more complex a grammar structure is, the more difficult it would be for L2 learners to learn. However, the fact that not all simple structures are easy to learn is a problem for this definition (Collins, Trofimovich, White, Cardoso, & Horst, 2009; Schmierer, 1979). For example, English simple past tense *–ed* and third person *–s* are simple in terms of their structure, but it takes time to use them spontaneously.

Larsen-Freeman (Celce-Murcia & Larsen-Freeman, 1999; Larsen-Freeman, 2003a, 2003b) discusses grammatical difficulty in terms of *linguistic form*, *semantic meaning*, and *pragmatic use*. "Linguistic form*"* concerns the accurate use of the (morpho)syntactic aspect of a grammar feature. "Semantic meaning" is primarily concerned with understanding the inherent or literal message encoded by a lexical item or a lexico-grammatical feature. "Pragmatic use" concerns using a lexico-grammatical feature appropriately within a context. According to Larsen-Freeman,

a grammar feature can be easy with respect to one aspect, but difficult with respect to another. For example, the form of the English passive is easy to learn, but its use is more difficult.

DeKeyser (2003) distinguishes *objective* difficulty from *subjective* difficulty. Objective difficulty concerns the linguistic factors that contribute to the learning difficulty of the structures in question. These factors include, for example, the complexity of the structure itself, the opacity of its meaning, and its scope and reliability. Subjective difficulty, on the other hand, takes individual learner differences into consideration. DeKeyser (2003) defines subjective difficulty as, "the ratio of the rule's inherent linguistic complexity to the student's ability to handle such a rule" (p. 331). In this regard, "what is a rule of moderate difficulty for one student may be easy for a student with more language learning aptitude or language learning experience" (p. 331). DeKeyser's view of subjective difficulty makes theoretical sense, and probably would be confirmed by observations of students' grammar performance.

R. Ellis (2006a, 2006b, 2008) considers grammatical difficulty in relation to implicit/explicit knowledge. Ellis (2006a) distinguishes two senses of *grammatical difficulty*: (1) "the difficulty learners have in understanding a grammatical feature," and (2) "the difficulty [learners] have in internalising a grammatical feature so that they are able to use it accurately in communication" (p. 88). Ellis argues that the first sense of grammatical difficulty relates to *explicit knowledge*, while the second sense relates to *implicit knowledge*. R. Ellis' (2006b, 2008) own studies support his argument that a grammar feature may be difficult to learn as explicit knowledge, but easy as implicit knowledge, and vice versa.

Collins et al. (2009) consider grammatical difficulty from an acquisition perspective. In their study that examined a corpus of aural input in relation to L2 learning, they define "easy rules" as those that tend to be acquired early, while the "hard rules" refer to those that tend to be acquired late. In line with this definition, English simple past, for example, is considered a difficult feature as it is acquired relatively late, while English progressives are easy features because they are usually acquired early. Collins et al. argue that viewing grammatical difficulty from the acquisition perspective has the advantage that it takes actual learner behaviour into consideration. However, the authors are also aware of two problems with this account. First, the equation of "early" to "easy" and "late" to "hard" suffers a problem of circularity; it does not

elucidate why a grammar form is easy or difficult to learn. Second, there is a lack of sufficient data to index early- versus late-acquired forms for all language features.

Scheffler (2009) determines grammatical difficulty from L2 learners' perspectives. In a study that explored rule difficulty in relation to the usefulness of L2 instruction, Scheffler used a questionnaire with a 5-point Likert scale to investigate 50 Polish EFL learners' perceptions of grammatical difficulty. The questionnaire asked the learners to rate the difficulty of 11 grammar structures from 1 to 5, with 1 indicating "very easy" and 5 indicating "very difficult." These 11 features were selected because the respondents became familiar with them in their classroom instruction at the time of the study. Results suggest that the learners felt that L2 instruction benefited them the most in learning the features they perceived to be difficult to learn.

Still other researchers have characterized grammatical difficulty in terms of students' correct use of the features (e.g., Ammar & Spada, 2006; Doughty & Varela, 1998; Spada, Lightbown, & White, 2005; J. Williams & Evans, 1998). Grammar features are considered more difficult to learn if many students have difficulty using them correctly. For example, the learning of the third person possessive determiner (his/her) is considered difficult for Francophone students learning English as an L2 because it has been frequently observed that the students tend to have difficulty using the feature correctly (Ammar, 2008; Lyster, 2004; Lyster & Izquierdo, 2009; J. White, 1998). Another example is English past tense *–ed*. It has been observed that accurate use of this feature is problematic for many L2 learners (R. Ellis, 2007; R. Ellis, Loewen, & Basturkmen, 2006; McDonough, 2007).

The foregoing discussion indicates that grammatical difficulty has been defined in a number of ways, including the extent of students' verbalization of the structure, comprehension versus production difficulty, the inherent complexity of the structure itself, ease or difficulty of learning form, meaning, and use, subjective versus objective difficulty, acquisition order, learners' perceptions of grammatical difficulty, and learners' use of grammar features. Such a multiplicity in the definitions of grammatical difficulty ironically bears witness to the fact that L2 grammar learning is complex!

To sum up, the question of what makes learning L2 grammar difficult is of considerable interest to L2 researchers and teachers. The SLA literature suggests that grammatical difficulty can be determined from psycholinguistic perspectives (e.g., developmental sequences) or

linguistic perspectives (e.g., inherent complexity of grammatical structures). However, while these theoretical accounts contribute to our understanding of the issue of grammatical difficulty, they are all "objective" accounts proposed by L2 theorists or researchers. Research on grammatical difficulty from the learners' perspective is still thin on the ground. Thus, insights into learners' perceptions of grammatical difficulty constitute a useful complement to theoretical accounts of grammatical difficulty. In addition, explorations of the relationship between learners' perceptions of grammatical difficulty and their L2 proficiency have the potential to inform L2 pedagogy in useful and important ways. Furthermore, the relationship between what features learners perceive to be difficult to learn and their ability to use those features is a question of both empirical and theoretical interest. The current study attempts to expand our understanding of these issues. Following Scheffler (2009), it investigates grammatical difficulty from the perspective of L2 learners. Specifically it investigates the issue of grammatical difficulty from the perspective of university-level Chinese EFL learners in relation to their overall L2 proficiency and L2 performance and knowledge.

Organization of the Thesis

The remaining chapters are organized as follows. Chapter Two provides a review of the literature on what constitutes grammatical difficulty, and discusses theories and research relating to second language proficiency, second language performance, and second language knowledge. The chapter ends with a discussion of the purposes of this research and the Research Questions. Chapter Three addresses the methods used in the current study and describes the research context, participants, research design, research instruments, and data collection procedures. Chapters Four and Five report on the data analyses and results for Research Questions 1 and 2, and 3 and 4, respectively. Chapter Six discusses the results relative to the four Research Questions, identifies limitations of the study, considers the implications of the findings, and provides suggestions for future research.

Chapter Two
Grammatical Difficulty, Second Language Proficiency, Performance, and Knowledge

This chapter consists of three sections. The first is a review of how the issue of grammatical difficulty has been addressed in the SLA literature. The second is a discussion of constructs and measurements of second language proficiency. The third considers theory and research related to second language knowledge and second language performance. The chapter concludes with a description of the purpose of this thesis research and the Research Questions.

A Review of the SLA Literature on Grammatical Difficulty

The question of what constitutes grammatical difficulty has long interested SLA researchers. In a paper that specifically addresses this intriguing question, DeKeyser (2005) ascribes grammatical difficulty to three factors: *problems of meaning*, *problems of form*, and *problems of form-meaning mapping*. According to DeKeyser, the meaning expressed through a grammatical form can be difficult to learn due to its "novelty, abstractness, or a combination of both" (p. 5). English articles is an example of a form that is difficult to learn due to its abstract meaning; learning English articles is particularly difficult for L2 learners whose L1s do not have an equivalent form. The form itself can also constitute a source of difficulty. For example, as DeKeyser notes, learners may have difficulty correctly using L2 morphological features because of, "the number of choices involved in picking all the right morphemes and allomorphs to express [these] meanings and putting them in the right place" (pp. 5-6). This is particularly true for richly inflected languages. In addition, a grammar form may be difficult to learn if (a) it contains non-essential communicative meaning, (b) its use is optional, and/or (c) its form-meaning relationship is obscure (for example, one meaning is expressed by multiple forms, or one form has multiple meanings). Learners learning a grammar form that fits any of these criteria may experience problems of form-meaning mapping. To illustrate, learning the English verb inflection *–ed* may be difficult because it appears redundant when the tense is also expressed by a temporal adverbial (as in a sentence such as *I walked to the park yesterday*). The optional use of case marking in Korean is also likely to result in difficulty establishing the form-meaning link. The morpheme *–s* in English represents a case of obscure form-meaning relationship; it can be

used to signify a third-person singular verb, a plural noun, or the genitive of a noun. Its obscure form-meaning relationship (or in DeKeyser's term, *opacity*) may cause learning difficulty.

Grammatical difficulty has also been discussed in the SLA literature with reference to several other factors, including:

(1) inherent complexity of rules,

(2) salience of a grammar form in the input,

(3) communicative force of a grammar form,

(4) input processing strategies in L2 learning,

(5) the L2 learner's developmental stage,

(6) L1 transfer, and

(7) individual differences in language aptitude.

In what follows, grammatical difficulty is discussed in relation to each of these seven factors.

Inherent complexity of rules.

It is frequently assumed that grammatical difficulty is in part dependent on the inherent complexity of *rules*; that is, the more complex the rules of a grammar form are, the more difficult it is for L2 learners to learn (Hulstijn, 1995). Although this assumption seems reasonable, there is no agreed upon standard for measuring the complexity of rules, making it difficult to characterize grammatical difficulty this way (de Graaff, 1997; Housen, Pierrard, & Van Daele, 2005). For example, there are disparities in definitions of the complexity of rules with respect to (1) what "rules" refers to, (2) how "complexity" is defined, and (3) how the complexity of rules is determined (Housen, et al., 2005). Regarding the first issue, two types of rules have been distinguished: linguistic rules (or "linguistic structure," in Housen et al.'s (2005) term) and pedagogical rules (Housen, et al., 2005; Robinson, 1996). *Linguistic rules* refer to, "the symbolic constructs postulated by linguists to denote or model observable linguistic phenomena (e.g., patterns of structural co-variance and form-function mappings) and/or their underlying mental representations" (Housen, et al., 2005, pp. 238-239). *Pedagogical rules*, on the other hand, refer

to, “a metalinguistic description of the explicit cognitive procedure” involved in producing a target linguistic rule (Housen, et al., 2005, p. 329). Pedagogical rules are instrumental in nature; they serve as tools for facilitating the learning of linguistic rules. It is this sense of rule that Housen et al. (2005) and Robinson (1996) have used in their investigations of complexity and L2 learning

Some researchers (Hulstijn & de Graaff, 1994; Robinson, 1996) distinguish complexity of rules from complexity of rule explanations. That is, linguistic rule complexity is the inherent complexity of the linguistic rules themselves, whereas pedagogical rule complexity is the complexities involved in explaining how a linguistic rule works (Housen, et al.; Robinson, 1996). Linguistic rule complexity can be further categorized into functional and formal complexity (DeKeyser, 1998; Doughty & Williams, 1998a). Simply put, functional complexity concerns form-function mappings, whereas formal complexity concerns the (morpho)syntactic constitution of a form. To illustrate, the use of English articles is *functionally* complex because this form has multiple functions. English present tense is *formally* simple because it is indicated by simply adding either ∅ or the morpheme *–s* to a verb’s base form.

With regard to the complexity of L2 rules, Hulstijn and de Graaff (1994) defined it as “the number (and/or the type) of criteria to be applied in order to arrive at the correct form” (p. 103). Following this definition, de Graaff (1997), in a study that investigated the effects of explicit instruction on L2 learning, operationalizes complexity as the total number of formal and functional grammatical criteria involved in the process of noticing, comprehending, or producing a given form. Basically, the fewer criteria required, the less complex the form. In a recent meta-analysis examining the effects of implicit and explicit instruction on the acquisition of simple and complex structures, Spada and Tomita (2010) also used the Hulstijn and de Graaff’s (1994) definition of complexity. Using their criteria, the authors characterized “wh-questions as object of a preposition” more complex than the simple past tense because the former requires seven transformations while the latter requires only one.

Housen et al. (2005) defined pedagogical complexity in a similar way, that is, in terms of “the number of steps the learner has to follow to arrive at the production of the intended linguistic structure, and the number of options and alternatives available at each step” (p. 241). In line with this definition, the researchers suggest that pedagogical rules for the formation of a

target structure can be more or less complex depending on the degree of elaboration with which the target structure is formulated. For example, the pedagogical rules for the formulation of the French present conditional can be as simple as "add the appropriate endings of the *imparfait* to the stem of the *future simple* form of the verb." The pedagogical rules in question can be complex if detailed information such as how to choose appropriate endings of the imparfait is provided.

House et al. (2005) investigated the effects of explicit instruction on L2 learning in relation to the issue of complexity. In their study, complexity is defined in terms of *functional markedness*, a concept advanced by Givon (1991, 1995). Givon's model of functional markedness comprises three major components: structural complexity, frequency and distribution, and psycho-cognitive complexity. According to Givon's model, one grammar form is considered to be more structurally complex than another if (1) producing the form requires more transformations of its underlying base form, (2) the form is not as frequently available to learners, (3) the use of the form is more strictly constrained by its syntactic and/or semantic context, and/or (4) acquisition of the form involves higher-level cognitive ability. Following Givon's criteria, Housen et al. concluded for the purpose of their study that the French passive voice is more complex than French sentence negation.

Recognizing the absence of agreed-upon, objective criteria for determining the complexity of rules, Robinson (1996) used teachers' judgments to establish rule complexity. Similarly, Van Baalen (1983, cited in de Graaff (1997), p. 251) used L2 textbooks and teachers' judgments to determine the complexity of rules.

It should be obvious from the previous discussion that various approaches have been employed to determine the complexity of rules. To make the complexity issue even more complex, even when researchers use the same definitions, they do not apply them in the same way. For instance, while Krashen (1982) and R. Ellis (1990) seem to have similar definitions of "formal complexity," Krashen classifies English third person simple present *–s* as formally simple because of the straightforwardness in term of describing its formulation, while R. Ellis considers it to be formally complex because of the distance between the verb stem and the noun phrase with which it agrees.

Although "complexity of rules" provides a seemingly compelling explanation for grammatical difficulty, given the variations in its definitions, the results of research that bases determinations of grammatical difficulty on complexity of rules should be taken with caution. In addition, it should be noted that language features with "simple" rules are not necessarily easily acquired features; English articles and third person *–s* are examples in point (Collins et al., 2009).

The salience of a grammar form in the input.

Grammatical difficulty has also been discussed in relation to the degree of *salience*[1] of forms in the input. Most SLA researchers agree that progress in L2 development requires attentional allocation to language forms in the input[2] (e.g., Doughty, 2001; Long, 1996, 2007; Lyster, 2007; Swain, 2005), although differences exist with regard to (1) the attention-awareness relationship (Gass, 1997; Robinson, 1995b, 2003; Schmidt, 1990, 1994, 1995, 2001; Skehan, 1998; Tomlin & Villa, 1994; VanPatten, 1996, 2004), and (2) the empirical findings regarding the necessity of awareness in L2 learning (Gass, Svetics, & Lemelin, 2003; Leow, 1997b, 2000; Rosa & Leow, 2004; Rosa & O'Neill, 1999). Based on the presumption that attention plays an essential role in L2 learning, L2 researchers argue that the more *salient* a form is, the more likely it is to be noticed and processed, and consequently acquired.

Salience of a grammar form is often discussed with reference to the "accessibility" and "availability" of the target form; the former is primarily contingent upon various linguistic attributes of the form, while the latter concerns the frequency of the form in the input to which learners are exposed (Collins, et al., 2009; Goldschneider & DeKeyser, 2005; Skehan, 1998). How easily a form can be perceived by the learner partially hinges on its linguistic attributes. In their meta-analysis of the determinants of the natural order of L2 morpheme acquisition, Goldschneider and DeKeyser (2005) posit that salience is determined by five components: *perceptual salience*, *semantic complexity*, *morphological regularity*, *syntactic category*, and *frequency*. A grammar form is more salient if the form is phonetically sonorous or stressed, semantically straightforward, morphologically predictable, and belongs to a syntactic category

[1] Here, salience is used in a broad sense rather than the narrow sense that considers only the attendant constraints pertinent to phonetic attributes in processing a form (Goldschneider & DeKeyser, 2005).
[2] The idea is not shared by researchers such as Krashen (1981, 1982, 1985, 1994) and Reber (1989), who claim that learning can be both conscious and unconscious, and that the latter is responsible for most L2 production.

that is more easily recognized.[3] Goldschneider and DeKeyser found that these five components collectively account for a considerable portion of the variance in predicting the order of acquisition of grammatical morphemes. Collins et al. (2009) analyzed a 110,000-word corpus of instructional talk to L2 learners, and found that the low phonetically perceptual salience of English simple past and the possessive determiners *his/her* partially explain why these two features are acquired relatively late by Francophone children learning English as L2.

Frequency of a grammar form also contributes to its salience (DeKeyser, 2005, Collins et al., 2009). The frequency-based account of SLA postulates that input frequency influences L2 learning. At a theoretical level, the frequency-based view of SLA is closely linked to the connectionist theory of language acquisition, which conceptualizes language knowledge as a network of interconnected nodes, positing that the creation and strengthening of the connections in the network largely depend on the learner's experiences with and sensitivity to input frequency (N. Ellis, 2002). In line with this view is the claim that the more frequently a form appears in the input, the more likely it is to be noticed, and thus acquired (Goldschneider & DeKeyser, 2005). However, researchers also note that for L2 learning to occur, input frequency operates in conjunction with other factors, such as the learner's L1, the learner's innate constraints on language learning, and linguistic attributes of the form in question (R. Ellis, 2009b; Gass & Mackey, 2002; Goldschneider & DeKeyser, 2005). That L2 learning depends on an interplay of a number of factors explains why there is a lack of an absolutely straightforward correspondence between input frequency and L2 learning. This is illustrated by the learning of English articles. English articles remain problematic for L2 learners even after massive exposure to the form (Master, 1994). Nonetheless, it still stands to reason that, at least for some grammar forms, the more frequently they appear in the input, the more likely they are to be noticed, and thus learned.

In addition to the accessibility and availability of a form in the input, salience is also attributed to factors such as the position of a language feature in a sentence or its communicative force (VanPatten, 2002). VanPatten and other researchers (e.g., Wong, 2004) argue that a grammar form positioned at the beginning or the end of a sentence gets more attention from the learner than one positioned in the middle of a sentence. In addition, a feature that communicates

[3] For example, lexical morphemes are more noticeable than functional morphemes.

essential information is more salient than a form conveying "redundant" information. (More detailed discussion of the communicative force of a grammar form is provided in the next section.)

Explaining grammatical difficulty with reference to the salience of a form in the input has the advantage of theoretical support (for example, from theoretical accounts of the role of attention in L2 learning or from connectionists' views of L2 learning). However, determining grammatical difficulty merely by salience is inadequate because, to date, we still lack a systematic understanding of, *inter alia*, what makes one grammar form more salient than another, and whether a grammar form is equally salient for L2 learners with different language proficiency levels, or for L2 learners with different L1s. In addition, as VanPatten (2007) reminds us, salience, "is often defined after the fact (e.g., something may not be salient if it is difficult to acquire)," and that it is, "vaguely defined to begin with and there is sometimes disagreement on just what the properties of salience are" (p. 178). In short, in light of these limitations, the degree of salience of a form is informative but inadequate for determining the difficulty of the form.

Communicative force of a grammar form.

Also related to the issue of salience and thus to the discussion of grammatical difficulty is the communicative force of a grammar form. According to VanPatten (2002), a form has communicative force if it contributes to the meaning of an utterance (VanPatten, 1996, 2007). In VanPatten's view (1996, 2004), the communicative force of a form depends on whether the form itself is semantically self-contained, and whether the form is semantically redundant[4] at the sentence level. In other words, a form is less communicatively valuable if its meaning can be retrieved from elsewhere in the sentence. However, a form may have different levels of semantic redundancy depending on the context in which it is used and the other forms it is used with (Harrington, 2004). Using VanPatten's criteria to illustrate, let us consider the sentence *It is raining*, which indicates that there is an event in progress by the progressive aspect marker *–ing*. Therefore, *–ing* is not a communicatively redundant feature in this sentence. In the sentence *I walked to work*, the verb inflection *–ed* is semantically important because it indicates the past

[4] Redundancy occurs when, "two or more elements in an utterance or discourse encode the same semantic information" (Farley, 2004, p. 7).

tense. However, when the past tense is captured by a temporal adverbial (for example, *I walked to work yesterday*), the verb inflection *–ed* is semantically redundant and thus less communicatively useful.

In terms of L2 learning, and arguing from an "input processing" perspective, VanPatten (2002) claims that a form with more communicative value is more likely to be noticed, and thus, "get processed and made available in the intake data for acquisition" (p. 760). This argument constitutes part of what VanPatten refers to as "the primacy of meaning principle," which posits that, due to processing capacity limitations, the learner's attention is prioritized toward semantic information before grammatical information while processing a linguistic input string.

Input processing strategies in L2 learning.

VanPatten (2002) argues that the use of inappropriate psycholinguistic processing strategies in L2 learning may also result in increased grammatical difficulty. The processing strategies he proposes are theoretically grounded in his model of "input processing (IP)." The IP model recognizes the crucial role played by "attention" in L2 learning, and maintains a single, limited capacity view of working memory.[5] The IP model postulates that a learner's processing capacity is limited during real-time comprehension so that he or she has to be selective in processing an input string, and that input processing taking place in the course of comprehension is likely to result in form-meaning mappings. Using the IP model, VanPatten argues that at the supra-sentential level, real-time processing demands may bias a learner to focus primarily on words that carry meaning at the expense of those with more abstract grammatical functions. And at the sentence level, the learner may be biased to interpret the first noun or pronoun as the subject of the utterance on his or her first encounter with the sentence.

In VanPatten's view, processing strategies can sometimes be sources of grammatical difficulty. VanPatten uses the learning of Spanish clitic object pronouns to illustrate his point. Spanish clitic object pronouns can take the initial position in a sentence. For instance,

La	*sigue*	*el*	*senor.*
Her-OBJ	follows	the	man-SUBJ.

[5] Working memory refers to, "the structures and processes that humans use to store and manipulate information" (Gass & Selinker, 2008, p. 250). Not all SLA researchers accept a single, limited capacity view of working memory. Robinson (2001), for example, postulates a multiple resources view of working memory.

This sentence translated into English is, “The man follows her.” Using an inappropriate word order processing strategy (in this case, the first noun principle), L2 Spanish learners, especially those whose L1s are rigid SVO languages, may misinterpret *La*, the object, as the subject of the sentence, and thus misunderstand the sentence as “She follows the man.” This Spanish example demonstrates that a word order processing strategy may lead the learner to make incorrect form-meaning connections.

Another example to illustrate the possible disadvantage of processing strategies for L2 learning is that of the English tense *–ed*. As previously illustrated, the verb inflection *–ed* is redundant in a sentence that also contains a temporal adverbial indicating past tense (for example, *I walked to work yesterday*). Therefore, using the “meaning first” processing strategy, the learner may overlook the verb inflection *–ed* in his or her processing of the input string, which could result in this grammatical feature being more difficult to learn. In short, grammatical difficulty may result from the processing strategies that are frequently, yet sometimes inappropriately, used by L2 learners in learning some grammar forms.

The IP model has made a considerable contribution to SLA theory by attempting to address the mechanisms that the learner uses to deal with novel input and the conditions favouring accurate input processing (Harrington, 2004). However, the IP model is not without its critics (DeKeyser, Salaberry, Robinson, & Harrington, 2002; Harrington, 2004). Major criticisms include, for example, a lack of definitional precision (i.e., multiple uses of the terms *meaning*, *form*, and *processing*) (Harrington, 2004) and a lack of clear accounts of what gets processed in the input (DeKeyser, et al., 2002). Accordingly, these caveats are useful to keep in mind in discussions of grammatical difficulty based on the IP model.

L2 learner’s developmental readiness.

The bulk of research on the acquisition of certain language structures suggests that many forms are learned in predictable stages. Evidence in support of staged acquisition of L2 grammar in English include findings obtained from studies of negation (Dulay, Burt, & Krashen, 1982), studies of pronouns (e.g., Zobl, 1985), studies of relative clauses (e.g., Pavesi, 1986), studies of possessive determiners (J. White, 1998; J. White, Munoz, & Collins, 2007; Zobl, 1985), studies of tense and aspect (e.g., Bardovi-Harlig, 2000; Shirai, 2004), and perhaps most compelling of

all, the work of Pienemann and his colleagues (e.g., Meisel, Clahsen, & Pienemann, 1981; Pienemann, 1989; Pienemann, Johnson, & Brindley, 1988) investigating the acquisition of word order in German and several morphological and syntactic features in English.

Staged L2 grammar development has been explained from various perspectives. For example, acquisition of tense and aspect has been explained from the perspective of the Aspect Hypothesis (e.g., R. Anderson & Shirai, 1996). The Aspect Hypothesis postulates that in the early stages of acquisition, the development of grammatical aspect markers is constrained by the temporal semantic meanings of verbs to which the markers are attached. The hypothesis has been supported by a number of empirical studies (see R. Anderson & Shirai, 1996; Bardovi-Harlig, 2000, for a review). Typological explanations such as the Noun Phrase Accessibility Hierarchy (NPAH) (Keenan & Comrie, 1977) have been employed to predict the acquisition orders of relative clauses. The NPAH suggests an implicational and unidirectional relationship among relative clause types. Based on the input frequency and the degree of "markedness"[6] of the relative clauses, the following hierarchical order of relative clauses has been mapped out: subject relative clauses (the most frequent and least marked type) > direct object relative clauses > indirect object relative clauses > object of preposition relative clauses > object of comparison relative clauses (the least frequent and most marked type). The NPAH has been supported by a number of studies of European languages (e.g., Doughty, 1991). However, studies of L2 Asian languages have not supported it, which makes its claims of universality less than convincing (Shirai & Ozeki, 2007).

In an attempt to explain a general pattern of L2 grammatical development in German and English, Pienemann (1998, 2003) proposed Processability Theory (PT). PT postulates that language learners travel across a series of predictable stages in their language development, and that learners' psycholinguistic processing abilities moderate their progress through the developmental stages. Drawing on Levelt's (1989) approach to language production, Pienemann (1998, pp. 73-86) proposes a hierarchy of five language generation stages: (1) word/lemma access, (2) category procedure (lexical category), (3) phrasal procedure (head), (4) S-procedure, (5) subordinate clause procedure. Pienemann argues that in the process of language production (acquisition), learners start from stage (1) and proceed sequentially. No skipping of stages is

[6] A "marked" linguistic structure is a structure that is "less basic," or "unique" with respect to others.

possible; each stage is a prerequisite for the next: access to higher-level procedures necessitates mastery of the procedures below. The processing sequences postulated by PT are strongly supported by Pienemann's own research in German, and moderately supported by research in other languages such as English, Scandinavian languages, Italian, and Japanese (Devitt, 2000).

L2 learners' "developmental readiness" is one criterion used by L2 researchers to determine grammatical difficulty. Adopting the view that the acquisition of certain grammar features takes place in predictable stages, some researchers have investigated the developmental sequences that L2 learners traverse in their learning of a certain target feature, for example, English question forms (e.g., Mackey, 1999; Mackey & Philp, 1998; Spada & Lightbown, 1993); relative clauses (e.g., Ammar & Lightbown, 2005; Y. Izumi & Izumi, 2004; Mellow, 2006); and English third person possessive determiner (e.g., Ammar, 2008; J. White, 1998; J. White, et al., 2007).

Explaining grammatical difficulty from the perspective of L2 learners' "developmental readiness" has theoretical advantages. However, unfortunately, to date, information about developmental sequences is only available for a handful of grammatical features in a small number of languages.

L1 transfer.

Discussion of grammatical difficulty resulting from "L1 transfer" can be traced back to the 1960s. Stockwell, Bowen, and Martin (1965) proposed the notion of "hierarchy of difficulty" in light of the Contrastive Analysis Hypothesis (CAH), which was originally advanced by Lado (1957). The CAH postulates that the degree of difficulty corresponds to the degree of difference between the target language and the learners' native language, and that the more differences there are between the two languages, the more difficult the target language will be for L2 learners. Thus, in Stockwell et al.'s view, grammatical difficulty is determined by L1-L2 differences. However, CAH is based on a behaviourist view of language acquisition and thus fell out of favour alongside behaviourism by the end of the 1970s, as did Stockwell et al.'s notion of "hierarchy of difficulty."

Language transfer has long been a focus of discussion in the literature (e.g., Gass & Selinker, 1992; Kellerman & Sharwood-Smith, 1986; Odlin, 1989, 2003; Selinker &

Lakshamanan, 1992), and the impact of L1 transfer on L2 learning has been a point of investigation in FFI studies (e.g., Izquierdo & Collins, 2008; Lightbown & Spada, 2000; Spada & Lightbown, 1999; L. White, 1991). Both theoretical accounts and empirical studies of language transfer indicate that L1 transfer may moderate the rate of L2 learning. Luk and Shirai (2009), in their recent review of English morpheme studies, further suggest a plausible L1 transfer effect on the developmental sequence of English morphemes. Accordingly, the authors argue that Krashen's (1982) "natural order" hypothesis should be used with caution due to the fact that it is primarily based on the analysis of Spanish learners' data.

"Negative transfer" resulting from L1-L2 differences may cause a certain degree of difficulty in L2 grammar learning. To illustrate, consider the problems, possibly derived from negative L1 transfer, that Chinese speakers have learning English *passive* and *conditionals*. I elect to discuss these two features because they are target features in the tasks conducted for the current study. The English passive is known to be problematic for Chinese ESL/EFL learners. Researchers argue that Chinese speakers' difficulty in learning this feature is, to some extent, attributable to L1 (Mandarin Chinese)-L2 (English) differences (Z. Han, 2000; Hinkel, 2002; Master, 1991). The differences between the Chinese passive and the English passive lie primarily in their linguistic properties and their frequency of use (Hinkel, 2002; Master, 1991; McEnery & Xiao, 2005; Yip, 1995). For example, in English, there are clear distinctions between active and passive constructions, whereas in Chinese, the distinctions are not clear, unless *bei* or one of its variants (e.g., *rang*, *jiao*) is present to indicate the passive. In addition, unlike the English passive, which is often used to indicate objectivity and a formal style, the Chinese passive often carries an adversative meaning. Furthermore, the English passive is used far more frequently than the Chinese passive. It is speculated that the less frequent use of the Chinese passive may be due to its adversative sense (McEnery & Xiao, 2005). The differences help explain why Chinese speakers have problems using the English passive construction correctly (Z. Han, 2000), or mastering English verb transitivity (Hinkel, 2002; Master, 1991), and why they tend to avoid using this feature (Hinkel, 2004).

Learning English conditionals is also quite challenging for Chinese ESL/EFL learners. One plausible reason for this is that English conditionals differ from Chinese conditionals in their linguistic representation (Chou, 2000). In English conditionals, the time of the event and the truth-value of the reference (or the degree of "hypotheticality") are primarily indicated through

the verb phrases (VP) in the main clause and in the *if*-clause (Celce-Murcia & Larsen-Freeman, 1999). However, whereas English conditionals primarily rely on syntactic features to indicate the time of the event and degree of hypotheticality, Chinese conditionals convey such information by the use of linguistic features at the syntactic, lexical and discoursal levels (Wu, 1994, cited in Chou, 2000). Chinese is a "tense-less" language; it uses special words to indicate the time of events. Wu argues that the counterfactual message in Chinese is often delivered by linguistic devices such as *temporal references*, *aspect markers* (e.g., the particle *le*), *negators*, and *rhetorical interrogatives.*[7] Li and Thompson (1981) suggest that a speaker engaged in a Chinese conversation usually infers the degree of hypotheticality from "the proposition in the second clause [of a conditional sentence], and from his/her knowledge of world, and of the context in which the sentence is being used" (Li & Thompson, 1981, p. 647). In terms of L2 learning, mastery of English conditionals not only requires sufficient, prerequisite knowledge of tense and aspect, modal auxiliaries, and negation (Celce-Murcia & Larsen-Freeman, 1999), but also a full understanding and good command of a variety of syntactic forms. Therefore, the learning of English conditionals is challenging for many L2 learners, perhaps particularly so for Chinese ESL/ESL learners, given the great differences between English and Chinese conditionals.

Individual differences in language aptitude.

It is assumed that grammatical difficulty is, to some extent, associated with individual differences in language aptitude. This is based on the view that learners with a stronger ability to learn languages may be better equipped to deal with grammatical difficulty in the L2 than learners with weaker aptitude. Language aptitude is the variable that contributes most to explanations of the considerable variation in "rates" and levels of attainment in L2 acquisition (Sawyer & Ranta, 2001). The conventional notion of aptitude was established by J. B. Carroll (Carroll & Sapon, 1959), who developed a multi-component model of aptitude and an attendant aptitude measurement battery, the *Modern Language Aptitude Test* (MLAT). The MLAT comprises five subtests designed to measure four underlying components of aptitude: phonemic coding ability, grammatical sensitivity, inductive language learning ability, and associative memory. The MLAT is by far the most widely used aptitude measurement battery. However,

[7] According to Wu (1994), the rhetorical interrogative is a unique characteristic of Chinese conditionals. Rhetorical questions are used in the main clause, and the if-clause is usually negative.

despite its popularity in research, criticisms against its validity have been documented, including, for example, that (1) its sub-tests mismatch with the components that they aim to measure, (2) it is based on an outdated account of memory capacity, and (3) it lacks an appropriate account of language processing (Sawyer & Ranta, 2001; Skehan, 2002).

In response to the criticisms of the conventional model of aptitude, revised models have been proposed. The models by Skehan (1998, 2002) and Robinson (2002) are particularly noteworthy. Skehan's (1998, 2002) model is characterized by an attempt to link components of aptitude to the L2 processing involved in L2 grammar learning. Skehan advances a reconceptualization of aptitude with three major components: auditory ability, language analytic ability, and memory capacity. Language analytic ability accommodates "grammatical sensitivity" and "inductive language-learning ability," which are two components of the MLAT. Considering L2 acquisition from an information processing perspective, Skehan maps out stages of L2 processing along with their attendant postulated processing demands and links them to specific aptitude components. Robinson's (2002b, 2005) model of aptitude includes not only the cognitive abilities underlying the aptitude construct, but also the possible interaction between aptitude and other variables such as task demands and learning conditions. Robinson (2002b), building on Snow's (1987, 1994) seminal work, employs the notion of "aptitude complexes" to explain the aptitude-treatment interaction. His main concern is to match, "learner's strengths in particular aptitude complexes to options in the delivery of learning conditions and instructional techniques at each of these levels" (Robinson, 2002b, p. 114). Both Skehan's and Robinson's models of aptitude suggest that aptitude is composed of multiple cognitive abilities. However, to date, not all the cognitive abilities included in the aptitude construct are well researched or well understood. Notwithstanding, these two models of aptitude provide us with a theoretical basis for the description of the "aptitude profiles" of L2 learners.

In spite of the criticisms of the conventional conception of aptitude, a number of studies have found it to be a useful predictor of language learning (de Graaff, 1997; Ehrman & Oxford, 1995; Erlam, 2005; Harley & Hart, 1997b, 2002; Robinson, 1995a; Sheen, 2007b). One inference that can be made from this finding is that grammatical difficulty is associated with individual differences in aptitude. Supporting evidence for this, although somewhat indirect, can be found in studies that suggest positive correlations between L2 learning and components of aptitude such as phonological short-term memory (Mackey, Philp, Egi, Fujii, & Tatsumi, 2002;

Robinson, 2002a; Trofimovich, Ammar, & Gatbonton, 2007) and analytic ability (Harley & Hart, 1997a; Ranta, 2002; Sheen, 2007a). However, because these studies are correlational in design, the findings do not allow us to make any causal inferences about grammatical difficulty and aptitude. Following Robinson (2002), a further inference that can be made is that grammatical difficulty may be attenuated if learners are provided with learning conditions that match their "aptitude profiles." The study by Wesche (1981) that investigated the effectiveness of matching aptitude with pedagogical methods provides some supporting evidence for this inference. Further supporting evidence comes from the study by Erlam (2005), which suggests that individual differences in language aptitude play a mediating role in determining the effects of the instructional methods under investigation.

Two caveats are in order with regard to the assumption that grammatical difficulty is related to individual differences in aptitude. First, as noted earlier, the aptitude construct is not yet fully understood. Second, the following questions have yet to be answered: whether aptitude matters more for adult L2 learners than for child L2 learners, whether aptitude makes a difference for L2 learners in different learning contexts (e.g., implicit, incidental, or explicit conditions), and whether the role of aptitude varies with the level of proficiency (Ortega, 2009). It goes without saying that more research into these areas would contribute to our understanding of the relationship between grammatical difficulty and individual differences in aptitude.

Summary.

The above discussion reveals that the question of what constitutes grammatical difficulty has been explored from various perspectives in the field of SLA, and that each of the accounts of grammatical difficulty has its own strengths and limitations. However, one message conveyed by these accounts is certain: L2 grammar learning is a complex phenomenon and more research on the issue of grammatical difficulty would contribute to our understanding of it.

Despite the availability of a multitude of accounts of grammatical difficulty, and regardless of whether they are linguistic or psycholinguistic in nature, they are all "objective" accounts proposed by L2 theorists or researchers. Little research has been conducted to investigate the issue of grammatical difficulty from the L2 learner's perspective. The current study is an attempt to fill this research gap by investigating the issue of grammatical difficulty

from the perspective of Chinese EFL learners. Specifically, it investigates which grammatical features Chinese EFL learners perceive as easier, and which as more difficult, to learn, and what influences their perceptions of grammatical difficulty. It also investigates the learners' perceptions of grammatical difficulty in relation to their overall L2 proficiency and their ability to accurately use grammatical features that they identify as more or less difficult to learn in the completion of different tasks.

Constructs and Measurements of Second Language Proficiency

One of the questions investigated in this study is the extent to which learners' perceptions of grammatical difficulty are related to their L2 proficiency. In this section of the literature review, I briefly consider how second language proficiency has been operationalized and measured in L2 research.

In language testing circles, views differ as to what constitutes the best model of L2 proficiency, although, currently, models based on a theory of communicative competence (Bachman, 1996) or on language performance (McNamara, 1996) seem to be preferred (R. Ellis, 2006b). In the field of SLA, however, L2 proficiency has been operationalized in a number of ways: for example, learners' overall L2 competence (Y. Han & Ellis, 1998; Harley & Hart, 1997a), learners' integrative ability to use a target language (Fotos, 1991; Laesch & van Kleeck, 1987), or learners' knowledge of grammar and vocabulary (Roehr, 2008b; Roehr & Ganem-Gutierrez, 2009).

There are also a number of different approaches to the assessment of L2 proficiency. Thomas (1994, 2006) surveyed how L2 proficiency has been operationalized and assessed in SLA research published in four key journals (*Applied Linguistics*, *Language Learning*, *Second Language Research*, and *Studies in Second Language Acquisition*) during two time periods (1988 to 1992 and 2000 to 2004). Based on her survey, Thomas reports that the proficiency-assessment techniques that are most common fall into four categories: *impressionistic judgment*, *institutional status*, *in-house measures*, and *standardized test scores*. According to Thomas, impressionistic judgment refers to unsubstantiated evaluations of learners' L2 competence (for example, learners' self-evaluation of their proficiency level or researchers' impressions of learners' proficiency); this is sometimes supplemented with information regarding the length of time that learners have spent living in a country where the target language is spoken. Institutional

status is determined identifying learners' as being members of a specific group (typically, an academic program). A statement such as "the participants were enrolled in the second semester of the German program at one university" exemplifies this proficiency-assessment technique. In-house assessment refers to either locally developed instruments (for example, program-internal placement tests) or tests designed for a specific research purpose and used to measure skills considered prerequisite to the target of investigation. Standardized tests refer to such as the Test of English as a Foreign Language (TOEFL) or the International English Language Testing System (IELTS), which are administered and scored in a consistent and standard manner.

The *cloze test* has also been used as a measure of L2 proficiency. A cloze test typically comprises a passage of approximately 300 words, from which 50 words have been deleted at regular intervals. The test requires that the test taker fill in each blank with the word that best fits. Scoring of the test is objective and can be done by the "exact-word" method or the "appropriate-word" method (Hanania & Shikhani, 1986): the former accepts only the words given in the prepared answer keys as correct answers, whereas the latter also counts as acceptable alternative correct answers. Research has shown that the results of using these two different scoring manners are highly correlated (e.g., Oller & Conrad, 1971). The cloze test is considered a measure of learners' integrative ability to use the target language (Fotos, 1991; Hanania & Shikhani, 1986). Hanania and Shikhani (1986) convincingly argue that:

> Cloze is considered an integrative rather than a discrete-point test because it draws at once on the overall grammatical, semantic, and rhetorical knowledge of the language. To reconstruct the textual message, students have to understand key ideas and perceive interrelationships within a stretch of continuous discourse, and they have to produce, rather than simply recognize, an appropriate word for each blank. The focus of the task involved is more communicative than formal in nature, and it is therefore considered to reflect a person's ability to function in the language. (p. 99)

Clearly, each L2 proficiency measure has its own advantages and limitations, and none of them is perfect. Discussion of the advantages and limitations of each L2 proficiency measure is beyond the scope of this review. However, one conclusion that can be drawn is that because there is so much latitude in the operationalization and measurement of L2 proficiency, these should receive careful attention in research in which L2 proficiency is an independent variable.

In this study, a cloze test developed by Fotos (1991) to assess learners' L2 proficiency was used. A rationale for the use of the test is provided in the Method chapter.

Second Language Knowledge and Second Language Performance

R. Ellis (2006a, 2006b, 2008) argues that grammatical difficulty differs with respect to type of L2 knowledge. An example to illustrate his point is that many L2 learners of English have no problem verbalizing the rule-of-thumb for the use of third person *–s*, but they have difficulty using it spontaneously.[8] Thus, this feature is considered to be easy to learn as explicit linguistic knowledge, but less easy to acquire as implicit linguistic knowledge. Researchers (e.g., DeKeyser, 2003; Doughty, 2003; Norris & Ortega, 2003; R. Ellis, 2005) suggest that L2 knowledge can only be inferred from learners' L2 performance, which can vary depending on the interplay of a multitude of task-related and learner-related factors. Motivated by R. Ellis's argument and considering the researchers' suggestions, the current study explored learners' perceptions of grammatical difficulty in relation to their L2 performance, and the findings are discussed largely with reference to type of L2 knowledge. Guided by this research purpose, this section of the literature review includes discussion of the following:

- Definitions of implicit and explicit knowledge
- Second language knowledge and second language learning
- The relationship between implicit and explicit knowledge and the role of explicit knowledge and explicit learning
- The use of second language knowledge in second language performance
- Second language knowledge and grammatical difficulty

Defining implicit and explicit knowledge.

In the field of SLA, there is a wide acceptance that *implicit* linguistic knowledge is distinct from *explicit* linguistic knowledge and that the development of implicit linguistic knowledge is the primary goal of L2 learning (R. Ellis, 2005). In the literature, various labels

[8] Early on, Krashen (1982) has advanced a similar idea. In line with his distinction between "acquisition" and "learning," Krashen argued that third person *–s* is easy to learn in terms of "learning," but difficult to learn in terms of "acquisition."

have been used to describe these two types of knowledge; for example, *acquired* versus *learned knowledge* (Krashen, 1981); *unanalyzed* versus *analyzed knowledge* (Bialystok, 1981; Bialystok & Ryan, 1985); *procedural* versus *declarative knowledge* (J. Anderson, 1983; DeKeyser, 1998); and *implicit* versus *explicit knowledge* (Bialystok, 1978; N. Ellis, 1994; R. Ellis, 2004; Hulstijn, 2005). Because the terms "implicit knowledge" and "explicit knowledge" are more often used in the SLA literature, they will be used here.

Despite the fact that cognitive psychology, developmental psycholinguistics and SLA provide a variety of definitions of implicit and explicit knowledge, these two types of knowledge are often distinguished with reference to (1) absence or presence of awareness/consciousness, and (2) ability (or the lack thereof) to verbalize linguistic knowledge on demand (e.g., J. Anderson, 2005; R. Ellis, 2005; Hulstijn, 2005; Roehr, 2006). In accordance with these two characteristics, R. Ellis (2009a, and elsewhere) argues that *implicit knowledge* can be characterized as "tacit and intuitive" and "only evident in learners' verbal behaviour," whereas *explicit knowledge* is "conscious" and "verbalizable" (R. Ellis, 2009a, pp. 11-13). In addition to these two characteristics, R. Ellis further differentiates these two by five other characteristics: *type of knowledge*, *accessibility*, *use of L2 knowledge*, *learnability*, and *systematicity*. R. Ellis argues that implicit knowledge is procedural knowledge (*type of knowledge*), accessible in automatic processing (*accessibility*) and employed in spontaneous L2 production (*use of L2 knowledge*). Learning of this knowledge may be affected by the age of the learner (*learnability*). In contrast, explicit knowledge is declarative knowledge and potentially available in controlled processing and for specific tasks (for example, completing a written grammaticality judgment test). Learning of this knowledge is less likely to be affected by the age of the learner. In addition, explicit knowledge is often inconsistent and inaccurate, whereas implicit knowledge is relatively systematic (*systematicity*). Moreover, like Krashen (1981, 1982), R. Ellis posits that explicit knowledge has greater utility when sufficient time is allowed and the focus of attention is on form. The characteristics reveal R. Ellis's attempt to conceptualize these two types of knowledge from various perspectives. These characteristics help us conceptualize the two types of knowledge, and from the perspective of testing, they are useful for the operationalization of the two types. However, as R. Ellis has acknowledged, some of these criteria are controversial and some are still subject to empirical verification.

R. Ellis (2004) further specifies the content of explicit knowledge. In his view, explicit knowledge comprises both "analyzed knowledge" and "metalanguage knowledge." In a subsequent article, he defines "analyzed knowledge" as "conscious representations of linguistic structures that can be verbalized on demand" (R. Ellis, 2006b, p. 437), and he defines "metalanguage knowledge" as learners' ability to use technical or semi-technical terminology to describe the language feature in use. R. Ellis (2004, 2006b; Y. Han & Ellis, 1998) contends that these two types of knowledge are distinct, yet the relationship between them is unclear.

Explicit knowledge developed by young adult or adult L2 learners primarily in formal learning contexts is sometimes referred to as "metalinguistic knowledge" (Alderson, Clapham, & Steel, 1997; Elder & Manwaring, 2004; Roehr, 2006). Metalinguistic knowledge is frequently associated with applied linguists' concept of *pedagogical grammar* (Roehr, 2008a, 2008b). Roehr (2008a) argues that metalinguistic knowledge is a specific type of explicit knowledge, and that it includes, "explicit knowledge about categories as well as explicit knowledge about relations between categories" (p. 72). Research on metalinguistic knowledge has revealed that instructed L2 learners vary considerably in the amount of the explicit knowledge they acquire (Alderson, et al., 1997; Green & Hecht, 1992), and their explicit knowledge often appears piecemeal, inaccurate, and idiosyncratic (Green & Hecht, 1992; Sorace, 1985). Berry (2005, 2009) refers to knowledge of metalinguistic terminology as "metalingual knowledge." Berry (2009) argues that "metalingual knowledge" should be distinguished from "metalinguistic knowledge (i.e., knowledge about language)" on the grounds that, "knowledge and the terminology for it do not always co-occur" (p.114).

Second language knowledge and second language learning.

Second language knowledge is closely related to second language learning, which is often discussed from the skill-acquisition perspective and the explicit-implicit learning distinction (Dornyei, 2009). The skill-acquisition accounts of L2 learning view language learning as analogous to the learning of a skill; in this view, L2 knowledge results from a process that progresses from acquired *declarative knowledge* to *procedural knowledge* through practice (DeKeyser, 1998, 2007). L2 declarative knowledge is learners' factual knowledge of the target language, while L2 procedural knowledge is their knowledge about how to perform the target language (DeKeyser, 1998).

From the explicit-implicit learning perspective, N. Ellis (1994) defines implicit and explicit learning as follows:

> Implicit learning is acquisition of knowledge about the underlying structure of a complex stimulus environment by a process which takes place naturally, simply and without conscious operations. Explicit learning is a more conscious operation where the individual makes and tests hypotheses in a search for structure. Knowledge attainment can thus take place implicitly (a nonconscious and automatic abstraction of the structural nature of the material arrived at from experience instances), explicitly through selective learning (the learner searching for information and building then testing hypotheses), or, because we can communicate using language, explicitly via given rules (assimilation of a rule following explicit instruction). (pp. 1-2)

N. Ellis's definitions indicate that the main difference between these two learning processes lies in the absence or presence of a conscious process of knowledge acquisition (for a comprehensive discussion of implicit and explicit learning, see DeKeyser (2003), N. Ellis (1994, 2005), Paradis (1994, 2009), and Schmidt (2001)). N. Ellis's view is broadly shared by other researchers (e.g., R. Ellis, 2005; R. Ellis, 2009a; Hulstijn, 2005; Schmidt, 2001) in their characterizations of these two types of learning processes. However, it should be noted that there is a lack of consensus with regard to the definition of "consciousness" (Robinson, 2003; Schmidt, 1990, 2001; Tomlin & Villa, 1994). In addition, there has been disagreement over whether any learning is possible without some degree of awareness/attention (e.g., Leow, 1997a; Schmidt, 1990, 1995, 2001; J. N. Williams, 1999, 2005). Despite these differences, however, two views are widely accepted: (1) implicit learning takes a relatively long time to be effective, and (2) explicit learning involves a certain degree of awareness/attention (Doughty, 2003; DeKeyser, 2003).

Although some researchers (e.g., J. Williams, 2005) have questioned whether explicit knowledge is inevitably the result of explicit learning, most researchers tend to agree that implicit and explicit knowledge result from implicit and explicit learning respectively (e.g., R. Ellis, 2004, 2005; Krashen, 1981; Schmidt, 1994). Empirical research has also shown a relationship between learning outcomes and learning conditions (e.g., Alderson, et al., 1997; Roehr, 2007; J. White & Ranta, 2002). For example, learners who learn an L2 primarily in grammar-oriented classes are observed to possess extensive explicit knowledge, but may be lacking in implicit knowledge. Despite the apparent relationship between implicit/explicit

learning and implicit/explicit knowledge, it would be unwise to assume that there is a straightforward causal relationship between learning conditions and learning outcomes (Bialystok, 1981). Besides, even assuming that there is a direct causal relationship, it does not follow that explicit learning or explicit knowledge plays no role in implicit L2 acquisition, a point that is discussed further below.

Relationship between implicit and explicit knowledge and the role of explicit knowledge and explicit learning.

The relationship between implicit and explicit knowledge and the role of explicit knowledge or explicit learning in implicit L2 acquisition have long been focal points of controversy. One point of debate is whether the implicit-explicit distinction describes a dichotomy or a continuum (R. Ellis, 2005, 2009a). Some researchers (N. Ellis, 2005; R. Ellis, 2004, 2005, 2009a; Hulstijn, 2002; Krashen, 1981; Paradis, 2009) view implicit and explicit knowledge as dichotomous, whereas others (Bialystok & Ryan, 1985; Dienes & Perner, 1999; Sharwood-Smith, 1981) take the continuum stance. Notwithstanding the theoretical disputes, neurological evidence has also indicated that these two knowledge types "involve different types of representation and are substantiated in different parts of the brain" (N. Ellis, 2005, p. 307), which provides some support for the dichotomy position (Paradis, 2009).

Another point of contention concerns the possibility of the interaction between the two types of knowledge at the *processing* or *learning* level, and a related point concerns the role of explicit knowledge or explicit learning in the development of implicit knowledge. Some researchers (Hulstijn, 2002; Krashen, 1981, 1982; Paradis, 1994, 2009; Schwartz, 1993) strongly reject any possible interaction between implicit and explicit learning, a position frequently referred to as the *non-interface position* (R. Ellis, 1994a, and elsewhere). Krashen (1981, 1982), Schwartz (1993), and Paradis (1994, 2009) are the most prominent defenders of this position. In an attempt to account for L2 acquisition, Krashen (1981, 1982, 1994; Krashen & Terrell, 1983) advanced the Monitor Theory, which has at its core a distinction between *acquired* knowledge and *learned* knowledge. According to Krashen, acquired knowledge is a consequence of a subconscious process that takes place when learners are using language for "real" communication in a positive affective state. Learned knowledge, on the other hand, results from the conscious process that goes on when learners turn their attention to information about

language such as that typically provided in grammar-oriented activities. Krashen terms the subconscious process *acquisition* and the conscious process *learning*. In his view, acquisition and learning are independent, and so are acquired and learned knowledge. Learned knowledge cannot be converted into acquired knowledge through practice and feedback. It can only function as a monitor that operates upon the output produced as a result of the acquired knowledge, and it is available for use only when learners know the rule, their attention is on form, and they have sufficient time. Krashen argues strongly that explicit knowledge plays a minimal role in the development of implicit knowledge, and is sceptical about the value of any form of formal instruction.

Schwartz (1993) distinguished *competence* (Chomsky, 1965) from what she called *learned linguistic knowledge*. The former is associated with Universal Grammar, whereas the latter, according to Schwartz, results from formal learning and negative evidence (i.e., information about ungrammaticality of an utterance). In her view, competence and learned linguistic knowledge are mutually exclusive; they "cannot communicate with one another" (Schwartz, 1993, p.160). In terms of the role of explicit learning, Schwartz argued that it does not contribute to the development of competence at the syntax aspect, but it may be beneficial for developing learners' competence of lexis and morphology.

Like Krashen and Schwartz, Paradis (2009) also argues strongly against the possibility of one type of linguistic knowledge being transformed into the other. However, unlike Krashen and Schwartz, whose arguments primarily rest on a Chomskian generativist view of language, Paradis relies heavily on neurological evidence to make his arguments. Drawing on this evidence, Paradis (1994, 2004, 2009) demonstrates that the two types of knowledge are sustained by separate memory systems; whereas implicit linguistic knowledge is sustained by procedural memory, explicit linguistic knowledge is sustained by declarative memory. These two memory systems subsist in separate, diffuse areas of the brain. In light of the fact that the two types of knowledge are different neurologically, Paradis argues emphatically against the possibility of a conversion of one type into the other through any form of learning or practice. However, although Paradis (2004, 2009) admits no role for explicit linguistic knowledge (or explicit learning) in the acquisition of implicit linguistic knowledge at the level of *learning*, he does acknowledge the usefulness of explicit L2 knowledge at the *production* level, claiming that it serves as a good resource that L2 learners may rely on when implicit linguistic knowledge does

not suffice for communication. In terms of the role of explicit learning, Paradis (2004) considers it to be beneficial in the sense that it can give rise to circumstances conducive to the attainment of implicit linguistic knowledge.

Contrary to the *non-interface* position, the *strong interface* position is based on the argument that one type of knowledge can be transformed into the other (R. Ellis, 2009b). One example of this position is Bialystok's (1978) early model of L2 learning.[9] In this model, Bialystok postulates that the development of different types of knowledge is primarily determined by the types of input to which L2 learners are exposed and the types of strategies that the learners use in their L2 learning. Explicit knowledge is developed through exposure to the input provided in form-oriented contexts and is promoted by the use of a strategy that focuses on the practicing of the language form itself (*formal practice*). In contrast, implicit knowledge is established through exposure to the input provided in meaning-based contexts and is facilitated by the use of a strategy that focuses on using the L2 for communication purposes (*functional practice*). Sufficient "formal practice" allows explicit knowledge to become implicit, and "inferencing" enables implicit knowledge to become explicit. Bialystok's (1978) position is shared by Sharwood Smith (1981), who also claims that explicit knowledge can be converted into implicit knowledge through practice.

The skill-acquisition accounts of L2 learning, which are based mainly on J. Anderson's (1983, 1990, 2005) Adaptive Control of Thought model (the ACT model),[10] also describe an "explicit-to-implicit" transformation (J. White & Ranta, 2002). J. Anderson's different versions of the ACT model are the most well-known models of skill acquisition. All the versions of the ACT model are grounded on a basic distinction between declarative knowledge and procedural knowledge. In addition, they all centre on a three-stage learning process: (1) a *cognitive* or *declarative* stage, where the learner consciously retrieves and encodes declarative information about the procedure stored in the declarative memory, (2) an associative stage, where the learner

[9] Bialystok's thinking about language learning has evolved considerably over the years. In her more recent model (the Analysis/Control model), the distinction between explicit and implicit knowledge has been replaced with two intersecting continua respectively manifesting knowledge types (analyzed versus unanalyzed) and processing processes (automatic versus controlled) (Bialystok, 1991, 1994).

[10] There are different versions of the ACT models, respectively referred to as ACT-E (J. Anderson, 1976), ACT* (J. Anderson, 1983), ACT-R (J. Anderson, 1993), ACT-R 4.0 (J. Anderson & Lebiere, 1998), ACT-R 5.0 (J. Anderson, Bothell, Douglass, Lebiere, & Qin, 2004). The general term "the ACT model" is used to refer to the model in all its versions.

acts upon the declarative knowledge and gradually finds it easier and easier to perform the procedures (i.e., a process called "*proceduralization*"), and (3) an automatic stage, where the performance of the procedures becomes automatic (for a detailed discussion, see J. Anderson (2000)). An essential component of the different versions of the ACT model is the proceduralization of declarative knowledge, which, according to J. Anderson's ACT theory (1993), is the gradual development of knowledge that is qualitatively different from the initially acquired declarative knowledge. In the learning process, learners will increasingly draw on this newly developed knowledge. However, J. Anderson's account of the qualitative conversion of knowledge is not without its critics (Dornyei, 2009; Segalowitz & Hulstijn, 2005; see also DeKeyser, 2001). Moreover, as J. Anderson et al. (2004) have acknowledged, the ACT model remains a work in progress and has limitations. Therefore, caution should be exercised when considering any strong claims based on the ACT model regarding implicit/explicit L2 learning and knowledge.

R. Ellis's and N. Ellis's positions regarding the role of explicit knowledge and explicit learning have been described as the "weak position" (R. Ellis, 2005 and elsewhere). R. Ellis (1994b) considers implicit and explicit knowledge to be "associable and cooperative." He postulates that the interface between the two kinds of knowledge can be "direct" or "indirect." It can be "direct" because explicit knowledge can be converted into implicit knowledge when learners are developmentally ready (Pienemann, 1989) to learn a target feature. It can be "indirect" because explicit knowledge can facilitate implicit L2 acquisition by promoting learners' awareness of the inadequacy of their linguistic repertoire (i.e., "*noticing a gap*," in Schmidt's (2001) term). N. Ellis (2005), on the other hand, considers the two types of knowledge to be "dissociable but cooperative" (p. 305); that is, knowledge conversion is unlikely, but explicit knowledge can *indirectly* contribute to implicit L2 acquisition. Based on connectionist accounts of L2 learning and drawing on a substantial body of research on language learning in relation to "frequency of linguistic input," N. Ellis (2002) views implicit learning and implicit knowledge as primary; however, he does not deny that "noticing," an internal cognitive mechanism proposed by Schmidt (1990, 2001), contributes to L2 learning. N. Ellis (2005) posits that explicit knowledge can facilitate implicit L2 acquisition in at least two ways: (1) by facilitating initial registration of a stimulus as a linguistic representation, which is consequently

accommodated in the implicit linguistic system, and (2) by serving as a processing constraint upon linguistic input received and then informing its interpretation.

In summary, the foregoing discussion indicates that there are divergent views with regard to the interaction between implicit and explicit learning. Regarding the role of explicit knowledge or explicit learning, unlike Krashen, who belittles the value of explicit knowledge and explicit learning in implicit L2 acquisition, other researchers tend to acknowledge a beneficial role for explicit knowledge or explicit learning. Nonetheless, disagreement as to how explicit knowledge/learning can contribute to implicit L2 acquisition persists. Regardless of the theoretical disputes, however, extensive relevant empirical research points to a facilitative role for explicit learning in implicit L2 acquisition, with supporting evidence from the findings of a substantial body of research on the roles of various forms of form-focused instruction (e.g., *corrective feedback*, *input enhancement*, *rule explanation*) in L2 learning (Doughty & Williams, 1998b; R. Ellis, 2001; Long, 1983; Norris & Ortega, 2000; Russell & Spada, 2006; Spada, 1997, 2010; Spada & Tomita, 2010; J. Williams, 2005). Nevertheless, the supporting evidence cannot all be taken without reservation given the methodological problems involved in measuring how language is learned under different conditions (Doughty, 2003; Norris & Ortega, 2003), a point to be further explicated and illustrated below.

Use of second language knowledge in second language performance.

As noted earlier, views differ considerably as to whether implicit knowledge and explicit knowledge interact at the *processing* or *learning* level. However, researchers tend to agree that L2 performance usually involves use of both implicit and explicit knowledge (DeKeyser, 2003; Doughty, 2003; N. Ellis, 2005; R. Ellis, 2005; Paradis, 2009). The extent to which implicit or explicit knowledge is brought into play may vary depending on the synergistic effect of a number of factors, such as task modality, time pressure, task requirements, learners' proficiency level, length and type of prior L2 study, to name a few (R. Ellis, 2005; Roehr, 2008b). In terms of task modality and time pressure, it is believed that, in general, an oral task under time pressure may encourage greater use of implicit knowledge, whereas a written task without time pressure may encourage more use of explicit knowledge (R. Ellis, 2005). However, it is also likely that learners adroit at accessing explicit knowledge may to some extent rely on it for real time oral communication (DeKeyser, 1998). Thus, neither task modality nor time pressure alone can

guarantee the exclusive use of implicit knowledge, although oral mode and time pressure may make the use of implicit knowledge more likely (DeKeyser, 2003; Norris & Ortega, 2003).

Apart from task modality and time pressure, task requirements also influence the degree to which implicit or explicit knowledge will be involved in the performance of a task (Bialystok & Ryan, 1985). To illustrate, what follows considers the potential utility of second language knowledge in oral production tasks and metalinguistic tasks, the two types of tasks that are frequently used in L2 research and in the present study.

Oral production tasks.

Oral production tasks are tasks that elicit the use of target features in spoken utterances. An oral production task that poses relatively few constraints and greatly encourages spontaneous, communicative use of target features in discourse (Norris & Ortega, 2000, 2003) is probably the best means to elicit the greatest use (but not exclusive use) of implicit knowledge (Doughty, 2003; Norris & Ortega, 2003). However, unfortunately, this type of task is difficult to design (Loschky & Bley-Vroman, 1993), and it may not allow "efficient" elicitation of target features. Thus, a more frequently used oral production task type is one that imposes certain constraints conducive to the elicitation of the target features. Oral production tasks of this kind may involve the use of explicit knowledge to varying degrees, depending on the task-related factors (Bialystok & Ryan, 1985). However, although research has revealed that learners' task performance can be influenced by a variety of factors such as participatory structure, direction of information (i.e., one-way to two-way), and topic familiarity (Yuan & Ellis, 2003), it remains unclear what factors tend to cause the use of explicit or implicit knowledge. Besides, it is also hard to disentangle the two types of knowledge or to measure the extent to which implicit or explicit knowledge has been used.

Metalinguistic tasks.

"Metalinguistic tasks" here broadly refers to any tasks that focus learners' attention on the language code itself in one way or another, including:

(1) judging overall grammaticality of a sentence;

(2) identifying an erroneous form of a sentence;

(3) correcting an erroneous form of a sentence;

(4) explaining why a form is erroneous;

(5) identifying named parts of speech in a sentence;

(6) any combination of (1) to (5).

Metalinguistic tasks using any combination of (1), (2), (3), and/or (4) have been referred to as *grammaticality judgments* (Bialystok, 1979; R. Ellis, 1991; Loewen, 2009), *acceptability judgments* (Bard, Robertson, & Sorace, 1996), *metalinguistic judgments* (Chaudron, 1983), and *error correction tests* (Spada, Jessop, Suzuki, Tomita, & Valio, in preparation). Metalinguistic tasks with a specific focus on (4) have been referred to as *tests of metalanguage* (R. Ellis, 2004), *metalinguistic tests* (R. Ellis, 2005, 2006) or *verbal reports* (Roehr, 2006).

Bialystok and Ryan (1985) postulate that metalinguistic tasks involving different ways of focusing on the language code itself may encourage access to and retrieval of what they called "analyzed knowledge," a concept somewhat equivalent to explicit knowledge. They argue that in terms of the analyzed knowledge brought into play, the tasks with the following requirements can be ranked (from moderate to high demand for analyzed knowledge) as follows: overall judgment of grammaticality of a sentence → identifying a deviant form of a sentence → correcting a deviant form → explaining why a form is deviant. However, views differ as to whether overall grammaticality judgment necessarily involves the use of explicit knowledge. Unlike Bialystok and Ryan, who suggest that overall grammaticality judgments involve a moderate degree of explicit knowledge, some researchers (e.g., Schachter & Yip, 1990) argue that the overall grammaticality judgment potentially reflects use of implicit knowledge, and thus can be taken as evidence of learners' linguistic competence. Nevertheless, the validity and reliability of using learners' grammaticality judgment performance as evidence of their L2 linguistic competence (or implicit linguistic knowledge) is not without critics (Bard, et al., 1996; Birdsong, 1989; Davies & Kaplan, 1998; R. Ellis, 1991; Goss, Zhang, & Lantolf, 1994; Y. Han & Ellis, 1998; Hedgcock, 1993). Notwithstanding these critiques, it has been reported that learners' overall grammaticality judgment performance can be influenced by factors such as "time span allowed for a response" and "task stimuli" (i.e., grammatical versus ungrammatical items) (Bialystok, 1979; R. Ellis, 2005; Loewen, 2009). Accordingly, R. Ellis (2005) and Loewen (2009) contend that *timed* judgments of *grammatical* items may predispose learners to

rely more on implicit knowledge, whereas *untimed* judgments of *ungrammatical* items may predispose learners to rely more on explicit knowledge.

According to Bialystok and Ryan (1985), tasks requiring learners to elaborate upon their linguistic knowledge predisposes them to rely to a great extent on analyzed knowledge. R. Ellis (2004) further argues that performing this type of metalinguistic task involves the use of both "analyzed knowledge" and "metalanguage knowledge." However, although there is general support for using learners' elaboration of their linguistic knowledge as evidence of explicit knowledge, it is also cautioned that this is by no means a sensitive or exhaustive measure of explicit knowledge as learners vary in their ability to verbalize (Bialystok & Ryan, 1985; R. Ellis, 1991, 2005; Y. Han & Ellis, 1998; Hu, 2002; Macaro & Masterman, 2006). Besides, research has also shown that learners can correct errors instantiating the rule in question even though they are unable to state the rule (Alderson, et al., 1997; Elder & Manwaring, 2004; Green & Hecht, 1992; Sorace, 1985).

The above discussion indicates that the presence or absence of time pressure and task modality (oral or written) in combination with specific task requirements may be conducive to the retrieval of one type of knowledge rather than another. It follows, then, that manipulating these factors may create conditions favourable for the demonstration of specific types of knowledge. Nonetheless, R. Ellis's (2005) word of caution is in order: learners are likely to employ, "whatever resources they have at their disposal irrespective of which resources are the ones suited to the task at hand" (Ellis, 2005, p. 153). R. Ellis's warning points to the essential role of learner-related factors in learners' task performance (see e.g., Robinson, 2001; Skehan, 1998; 2001, for discussions). Of various learner-related factors, learners' current L2 knowledge is one that may influence their task performance. It is likely that learners well equipped with the type of knowledge that meets the task requirement may excel in performing that kind of task, but they may not do as well on tasks that require the other type of knowledge. To illustrate, learners with extensive explicit knowledge are very likely to perform well on metalinguistic tasks. However, learners with inadequate implicit knowledge may not perform well on tasks that require extensive implicit knowledge. In short, learners with different repertoires of explicit or implicit knowledge may perform differently on tasks that make different demands on their L2 knowledge.

It should be obvious from the foregoing discussion that measuring L2 knowledge, particularly implicit knowledge, is indeed a great challenge for L2 researchers. The crux of the challenge lies in the fact that the presence of the implicit is inferred rather than directly observed, thus the reliance on "behavioural" evidence of implicit knowledge is inevitable (DeKeyser, 2003; Hulstijn, 2002). Besides, learners' L2 performance may engage implicit and explicit knowledge to varying extents depending on an interplay of a number of factors. Despite these methodological challenges, perhaps we can consider DeKeyser's (2003) suggestion that, "researchers have to content themselves with eliciting knowledge under conditions that are more or less conducive to the retrieval of implicit and explicit knowledge, and then infer to what extent the learning itself may have been implicit or explicit" (p. 320). In addition, as every test or task intended as a measure of implicit or explicit knowledge has design limitations and/or its results may be subject to different interpretations (Chaudron, 2003), the use of multiple tests is preferred because it allows for the collection of various kinds of performance data, and thus provides a more complete picture of L2 learners' linguistic knowledge (Norris & Ortega, 2003).

Second language knowledge and grammatical difficulty.

In the SLA literature, R. Ellis (2006b, 2008) is one of the few who has addressed the issue of grammatical difficulty in relation to implicit and explicit knowledge. He distinguishes two senses of *grammatical difficulty*: (1) "the difficulty learners have in understanding a grammatical feature," and (2) "the difficulty [learners] have in internalising a grammatical feature so that they are able to use it accurately in communication" (R. Ellis, 2006a, p. 88). He argues that the first sense of grammatical difficulty relates to explicit knowledge, while the second sense relates to implicit knowledge. R. Ellis (2006b) postulates that the following criteria might function as determinants of grammatical difficulty at the level of implicit knowledge: (1) *frequency* (i.e., whether the grammar features occur frequently in the input); (2) *salience* (i.e., whether they are relatively noticeable in the input); (3) *functional value* (i.e., whether they have essential communicative force); (4) *regularity* (i.e., whether they are regular or irregular grammar features); and (5) *processability* (i.e., whether they are easy to process based on Pienemann's (1998) processing theory). Despite the availability of theoretical support for these criteria, R. Ellis (2006b) acknowledges that, "it is not clear how such criteria can be applied to determine the learning difficulty of different grammatical features" (p. 437).

With regard to grammatical difficulty at the level of explicit knowledge, R. Ellis (2006b) proposes the following, all of which, as he acknowledges, are in need of empirical verification.

- Grammar features that are formally and/or functionally simple are easier to learn than those that are formally and/or functionally complex.
- "Rule-based" grammar features are easier to learn than "item-based" grammar features.
- Grammar features with simple pedagogical rules (e.g., the use of third person –*s*) are easier to learn than those with relative complex pedagogical rules (e.g., the choice of articles).
- Grammar rules specifying the prototypical function of a form are easier to learn than those specifying the peripheral function of the form.
- Grammar rules that can be formulated with less extensive use of metalanguage are easier to learn than those that need more extensive use of metalanguage.

To investigate whether grammatical difficulty varies in terms of type of L2 knowledge, R. Ellis (2006b, 2008) conducted two studies, both of which used the same measures of implicit knowledge and explicit knowledge. Explicit knowledge was measured by an untimed grammaticality judgment test and a metalinguistic knowledge test, whereas implicit knowledge was measured by an elicited imitation test, an oral narrative test, and a timed grammaticality judgment test. The first study (R. Ellis, 2006b) targeted 17 features. The learning difficulty of these features was determined by the accuracy of the participants' performance of them on the tests. Results showed that the difficulty order of the 17 target features resulting from the tests measuring explicit knowledge differed from that resulting from the tests measuring implicit knowledge. In other words, some features observed to be difficult in terms of implicit knowledge were not difficult in terms of explicit knowledge, and vice versa. However, the findings should be taken with caution since, as R. Ellis himself acknowledges, the number of test items for each individual feature is relatively small.

The second study (R. Ellis, 2008) targeted only four features, grammatical difficulty of which was determined by the use of Pienemann's (2005) Processability Theory. The four target features were possessive -*s*, since/for, 3rd person -*s*, and question tags; they were listed in the order of their difficulty level (from the least to the most difficult level). The predicted difficulty

order of the four target features was borne out in the tests measuring implicit knowledge, but not in the tests measuring explicit knowledge. In light of the findings, R. Ellis concludes that grammatical difficulty is relative to implicit and explicit knowledge. However, as only a small number of the test items were used for each feature, the findings are at best suggestive.

Summary.

Despite widespread acceptance that implicit knowledge is distinct from explicit knowledge and that the development of the former is the primary goal of L2 learning, views differ considerably with regard to the relationship between the two types of knowledge and the role of explicit learning or explicit knowledge in implicit L2 learning at the processing or learning level. However, at the L2 performance level, researchers tend to agree that the combined use of implicit and explicit knowledge is involved. Given that L2 knowledge can only be inferred from L2 performance, which can vary depending on a multitude of task-related and learner-related factors, measuring L2 knowledge presents a great challenge for L2 researchers. Research on the relationship between L2 knowledge and L2 proficiency indicates the two are correlated. However, the findings should be taken with caution given the differences in the operationalizations and assessments of L2 knowledge and L2 proficiency. R. Ellis's (2006b) claim that grammatical difficulty is relative to implicit and explicit knowledge is theoretically convincing. Nevertheless, how grammatical difficulty relates to these two types of knowledge remains unclear since the relationship between grammatical difficulty and type of L2 knowledge is still under-researched.

Research Purposes and Research Questions

The above review of the literature indicates that further empirical research on grammatical difficulty is desirable. This study investigates grammatical difficulty from the learners' perspective. Furthermore, it explores whether and how learners' perceptions of grammatical difficulty relate to their overall L2 proficiency, L2 knowledge and performance. The study aims to explore the issue of grammatical difficulty from the perspective of university-level Chinese EFL learners. The specific questions motivating this research are:

1. Which features of English grammar do university-level Chinese EFL learners perceive as easier, and which as more difficult, to learn?

2. What influences Chinese EFL learners' perceptions of grammatical difficulty?
3. Do Chinese EFL learners' perceptions of grammatical difficulty vary according to their English proficiency level? If so, how?
4. How do Chinese EFL learners' perceptions of grammatical difficulty relate to their ability to accurately use grammar features on oral production and written metalinguistic tasks?

Chapter Three
Method

This chapter starts with a description of the research context, and continues with an overview of the study outlining its three distinct phases. This is followed by a description of the research design and methods employed in each phase.

Research Context – EFL Teaching and Learning in Taiwan

The study was conducted in two national universities in the central part of Taiwan. In Taiwan, EFL teaching at the *secondary* school level is primarily form-oriented despite the fact that the curriculum guidelines for high school EFL instruction clearly specify that the main goal is to help students develop communicative ability. The EFL teaching and learning reality at the secondary school level is consistent with Fotos's (2005) observation of English learning in some EFL contexts:

> [Many] EFL situations have a centrally controlled education system with a set of curriculum, prescribed textbooks, and highly competitive nationwide examinations determining admission to middle, second, and tertiary institutions. Such examinations usually have an English component requiring reading comprehension, knowledge of grammar rules, vocabulary, and translation skills. As a result, English language teaching is often aimed at mastery of the points tested on such examinations . . . .EFL instruction usually emphasizes the development of knowledge about English, including grammar rules, and the development of vocabulary and translation skill, rather than the development of communicative ability. (Fotos, 2005, pp. 665-666)

EFL teaching and learning at the *university* level is primarily skills-based. First-year undergraduates take reading as well as listening-and-speaking courses, though English writing is often excluded from the EFL curriculum due to practical constraints (e.g., inadequate numbers of EFL teachers). In recent years, with an eye to improving the quality of tertiary education, the Ministry of Education (MOE) in Taiwan has been encouraging universities to create "Teaching Excellence Programs (TEP)," a set of schemata to improve teaching in different academic fields. Although the focus of TEPs varies from one university to another, EFL instruction is often an important aspect. As a result, many universities have taken measures to enhance their students' English learning, including offering a greater variety of credited EFL courses and non-credited English learning activities, providing more teaching and self-study materials, establishing on-line

learning platforms, and requiring students to pass an exit exam in EFL in order to graduate.[11] In short, in Taiwan, recent years have seen increasing attention paid to EFL teaching and learning in tertiary education.

The study was conducted in two national universities with funded TEPs that give prominence to EFL instruction. The universities provide their students with abundant English learning resources, and include an English exit exam as part of the requirement for graduation. The universities require all first year students to take a credited English reading class, but specify the listening-and-speaking class as optional. Therefore, the weekly in-class EFL instruction time for the first year students ranges from two to four hours–two hours if they just take the required course and four hours if they take the optional course. There are no prescribed textbooks or syllabi for any EFL courses in these two universities: EFL instructors are free to choose the textbooks that they think are appropriate for their students and to design their own syllabi. One university requires their first year students to participate in an on-line English learning program, designed by a team of EFL instructors. The on-line learning emphasizes English vocabulary, grammar, and phrases, and is monitored by EFL instructors.

The study involved four English instructors and students in seven of their classes. At the time of data collection, grammar was not the main focus of their instruction; however, two teachers reported that they sometimes reviewed grammatical features that appeared to be problematic for their students.

Overview of the Study

The study employed multiple research methods to address the Research Questions in order to "triangulate" the findings and to offset the disadvantages of using a single method. Figure 1 provides an overview of the study's three distinct phases. Phase 1 involved the administration of a questionnaire and a proficiency test to investigate Research Questions 1 and 3. Phase 2 comprised a semi-structured interview, a grammatical difficulty ranking activity, two cloze activities, and two stimulated recall activities, all of which were intended to explore

[11] In order to graduate, undergraduates need to pass one of the following English proficiency tests: TOEFL, TOEIC, IELTS, or GEPT (General English Proficiency Test, a locally designed English proficiency test), or another English proficiency test recognized by the universities.

Research Questions 1 and 2. Phase 3 involved the administration of oral production tasks and a written metalinguistic task to explore Research Question 4.

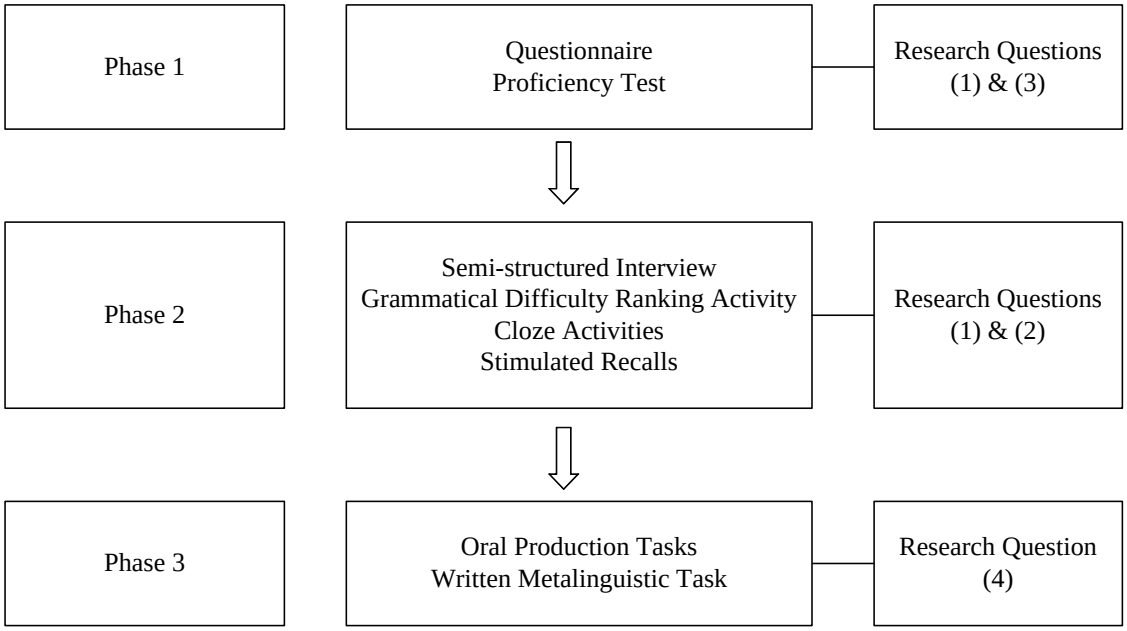


***Figure 1*. Overview of the study.**

Phase 1

Participants.

In Phase 1, a questionnaire and a proficiency test were administered to a group of 277 university-level Chinese EFL learners. The participants were drawn from seven intact classes, including five English reading classes and two credited English listening-and-speaking classes. The reading and listening-and-speaking classes were primarily intended for first year students; however, senior students who had previously failed these two courses could take them as well. The selection of the classes was based mainly on the instructors' willingness to have their class participate. Four English instructors were interested; thus, the students in the seven classes taught by these four instructors were invited to participate in this study. More details concerning the recruitment of the teacher and student participants will be discussed in "Data Collection Procedures."

To maximize the homogeneity of the sample, out of the 293 students enrolled in the seven classes, 16 were excluded. Participants were excluded if (1) they were non-first-year students, (2) they were international students, and/or (3) they had spent more than one year in an English-speaking country. As a result, the number of participants for the questionnaire and the proficiency test data analysis was reduced to 277. These 277 participants were more or less equally distributed among the seven classes: 30 in Class 1,[12] 46 in Class 2, 43 in Class 3, 29 in Class 4, 39 in Class 5, 45 in Class 6, and 45 in Class 7. The participants were from the following academic majors: humanities (43.3%, $n = 120$), social science (21.7%, $n = 69$), and science (35%, $n = 97$).

The participants were relatively homogeneous in terms of their cultural background, native language, and English learning history. Among the 277 participants, 116 were male and 161 were female. Mandarin Chinese was the first language of 231 participants, 39 spoke Taiwanese, and 7 spoke other languages as their mother tongue. The average age of the participants was 18 years, ranging from 17 to 20. At the time of data collection, the participants had studied English for 8.8 years on average, with the range between 5 and 13 years ($SD = 2.16$). The majority of the participants reported that they had received several years of grammar-based instruction in high school. They also reported that, in general, their informal exposure to English was limited to listening to English songs, watching English movies, and reading English magazines.

Research instruments.

Questionnaire.

In order to obtain a general understanding of Chinese EFL learners' perceptions of the difficulty of learning various English grammatical structures, I designed and administered a questionnaire. I used the questionnaire format because it has the advantage of gathering a large amount of information within a short time and provides results that are easily quantified and analyzed (Dornyei, 2003; Gillham, 2007).

[12] Class 1 participated in the pilot study; however, because no changes were made to the instruments or administration of tasks based on the pilot, it was decided to include these data in the main study.

The questionnaire consists of three sections (see Appendix A). Section 1 asks for the participants' biographical information, including sex, age, prior EFL learning experience, informal exposure to English, and English proficiency test scores. Section 2 comprises 20 closed-ended questions, each of which represents a different grammatical feature. The selection of the target features was based on three criteria: (1) they are covered in the high school teaching syllabus, (2) they are known to be problematic for Chinese EFL learners, and (3) they are morphological and/or syntactical in nature. Using these criteria, I reviewed a number of English textbooks and grammar books used in high school in Taiwan. I also consulted with several high school English teachers. Eventually, the following features were selected:

- present perfect
- simple past
- negation
- modal auxiliaries
- countable/uncountable nouns
- passives
- articles
- unreal conditionals
- embedded questions
- third person *–s*
- clauses
- present progressive
- prepositions
- adjective comparatives
- past progressive
- infinitive
- wh-questions
- participial construction
- question tags
- real conditionals

In the questionnaire, parts of speech are used to describe each grammar structure. To help the participants understand the parts of speech, two sample sentences using the target structure were provided, with the target structure underlined (for example, *I <u>have finished</u> the job*.). In Section 2, the participants were asked to indicate the degree of difficulty using a six-point Likert scale, with 1 standing for "Not at all difficult" and 6 standing for "Extremely difficult." "Not at all difficult" indicates that the student had learned the structure quickly after a short explanation and some practice. "Extremely difficult" indicates that the student does not expect to ever understand the structure fully, even with extensive explanation and practice. The participants were asked to base their rating on their prior grammar learning experience. In Section 2, an even-

numbered Likert scale (6-point scale) was used in order to avoid the "neutral" responses that are frequently obtained in an odd-numbered (e.g., 5-point scale) Likert scale (Brown, 2001; Dornyei, 2003).

Section 3 (the qualitative section) comprises four reflective items, that ask the participants to explain their assessment of the difficulty of the selected features. Four features of English grammar (third person -*s*, passives, articles, and unreal conditionals) were selected for this section, which was designed to explore possible reasons for the ease or difficulty of learning these four features from Chinese EFL learners' perspectives. These four features were selected because they have been observed to be problematic for Chinese learners to use based on my teaching experience and that of my colleagues. I also wanted to select features that are considered difficult for these learners even though some are taught earlier in their EFL instruction (e.g., third person –*s* and articles) than others (e.g., passives and unreal conditionals).

To help the participants better understand the questionnaire, a Chinese-English bilingual version of it was provided. The participants were invited to answer the reflective questions in any language they felt comfortable using.

The questionnaire was piloted in one of the seven classes that participated in the questionnaire survey (i.e., Class 1). On average, it took the students in this class 30 minutes to complete it, with a range from 20 minutes to 40 minutes. Of the 31 students who completed the questionnaire in the pilot study, 1 did not meet the selection criteria, and thus was excluded from the subsequent data analysis (see the section "Participants" for the criteria). The data from the other 30 participants were analyzed using the Statistical Package for the Social Science (SPSS) for Windows 15.0. To assess the reliability of the questionnaire, I used Cronbach's alpha index. The index is widely used to assess the degree of internal consistency or reliability of instrument used in data collection (Brown, 2001). The results showed that the questionnaire has a high degree of reliability, Cronbach's alpha .91. In addition, it was observed that the participants had a good understanding of the metalinguistic terminology used. Given the high reliability and comprehensibility of the questionnaire, no changes were made in the closed-ended question section. The reflective items also remained the same in light of the observation that the items seemed to be comprehensible to the students. Given the pilot study findings, the questionnaire unaltered was given to the other six classes.

Proficiency test.

To measure the participating learners' overall English proficiency, I used a cloze test, developed by Fotos (1991) (see Appendix B).[13] Fotos's cloze test was adapted from an intermediate level ESL book, with a 1,100- to 1,500-word range. The test consists of three paragraphs, 399 words, with seventh-word deletions. The deleted words, 50 systematically spaced blanks, fulfill a variety of grammatical functions.

I used Fotos's cloze test for the following reasons. First, Fotos (1991) makes a compelling case that the cloze test is a reliable indicator of EFL proficiency. She does so by demonstrating that students' cloze test performance significantly correlated with their performance on an essay test designed to measure their integrative language ability. Learners' scores on the cloze also significantly correlated with their performance on the TOEFL. Second, the participants in my study were familiar with the cloze test format, as this kind of test is frequently used in high schools in Taiwan. Thus, the influence of lack of test familiarity is of little concern. Third, the cloze test takes much less time than other proficiency measures such as TOEFL or IELTS. It can be completed within 45 minutes compared with the several hours required for other general measures of proficiency. Although the use of cloze tests as measures of overall proficiency is not without its critics (see, for example, Brown, 1983, 1989; Hanania & Shikhani, 1986), considering its time effectiveness and familiarity to students, it was decided to use the cloze test for this purpose.

The cloze test was also piloted in Class 1. Following Fotos (1991), the students were allowed a maximum of 45 minutes to take the test. The observation made on-site indicated that 45 minutes was reasonable for the students to complete the test. I assessed the reliability of the test by using Cronbach's alpha index, which indicated a high level of reliability (.85). Thus, the cloze test was administered to the other classes unchanged. Because no changes were made to the test based on the pilot, the data from the pilot class were included in the main study.

---

[13] From "The cloze test as an integrative measure of EFL proficiency: A substitute for essays on college entrance examinations?" by S. Fotos, 1991, *Language Learning*, 41, p. 336. Copyright (2010) by the name of copyright holder. Reprinted with permission.

Data collection procedures.

Phase 1 research was conducted in the second week of December 2008. Prior to conducting the study and after obtaining ethics approval for the research, I sent a letter of inquiry[14] to department heads to ask for permission to contact EFL instructors in their departments. With the department heads' permission, I then contacted several EFL instructors through emails to see if they were interested in the study. Four teachers showed interest, and volunteered seven of their classes for the study. One week before Phase 1 was conducted, I went to "Class 1" (i.e., the pilot class) to pilot the questionnaire and the proficiency test. After the pilot, with teachers' permission and at the classes' convenience, I went to the other six classes to administer the questionnaire and the proficiency test.

The pilot and the main study followed similar data collection procedures. I first explained the study, and distributed the Chinese version of the letter of inquiry for students (see Appendix C). I then distributed the Chinese version of the consent form[15] (see Appendix D) to the students. After the signed consent forms were collected, I administered the questionnaire. Students were allowed a maximum of 45 minutes to complete the questionnaire. When 45 minutes were up, I collected the questionnaires from the students and gave them a 10-minute break. After the break, the students completed the proficiency test. They were allowed a maximum of 45 minutes and the tests were collected when the time was up.

Phase 2

Participants.

Phase 2[16] involved 30 of the participants from Phase 1. In this phase, all 30 participants met with me individually and completed the following activities: a semi-structured interview, 2 cloze activities, and 2 stimulated recall activities. These 30 participants were selected from a pool of volunteers who had participated in Phase 1. During Phase 1, I asked the participants to

[14] There are three versions of the letter of inquiry, one for each of the department heads, the EFL teachers, and the prospective student participants. As the letters are very similar in content, only the Letter of Inquiry for students is included in the Appendices.

[15] There are three versions of the consent form, one for each of the department heads, the teachers, and the students. Because the three forms are similar in content, only the consent form for the students is included in the Appendices.

[16] In the letter of inquiry and the consent form, Phase 2 research was referred to as "semi-structured interview" for the sake of convenience and simplicity. However, the various activities included in the "interview" were described in further detail in the letters and the consent forms.

indicate at the end of the questionnaire whether they were interested in participating in Phase 2 of the research. Those interested were asked to leave their contact information and to indicate the best time slots for them to meet with me individually. In total, around 90 students showed interest; from among them, I randomly selected 40 students. In the end, 30 out of the 40 students participated in this phase. Of these 30 participants, 11 were male and 19 were female. More details about these 30 students are provided in Chapter Four. To compensate for their time, each participant was given a gift coupon worth the equivalent of 3.3 Canadian dollars (100 New Taiwan dollars).

Research instruments.

Semi-structured interview.

A semi-structured one-on-one interview was conducted with each of the Phase 2 participants to obtain background information and to serve as a warm-up for the subsequent activities (see Appendix E for the interview protocol). During the interview, I asked each participant questions about his or her prior experience with English grammar, enjoyment or not of grammar learning, and attitudes toward the usefulness of grammar lessons. The interviews were all conducted in Mandarin Chinese. With the participants' permission, all of the interviews were audio-recorded.

Grammatical difficulty ranking activity.

Out of the 20 grammar features in the questionnaire, I selected 10 for the grammatical difficulty ranking activity (see Figure 2 below). In this activity, the participants were asked to rank the 10 features for learning difficulty and then to explain the rationale behind their ranking. The ranking results obtained from this activity were used to verify the ranking results obtained from the questionnaire survey. Only 10 of the 20 features were used because this was considered to be a more practical number for this kind of activity. The difficulty ranking (from 1-20) for the 20 features used in the questionnaire survey appears in the left hand box in Figure 2. The 10 features marked with an asterisk (*) are the ones selected for the ranking activity. While the selection of these 10 features was rather arbitrary, they are more or less evenly distributed in terms of their difficulty level, as revealed by the questionnaire findings.

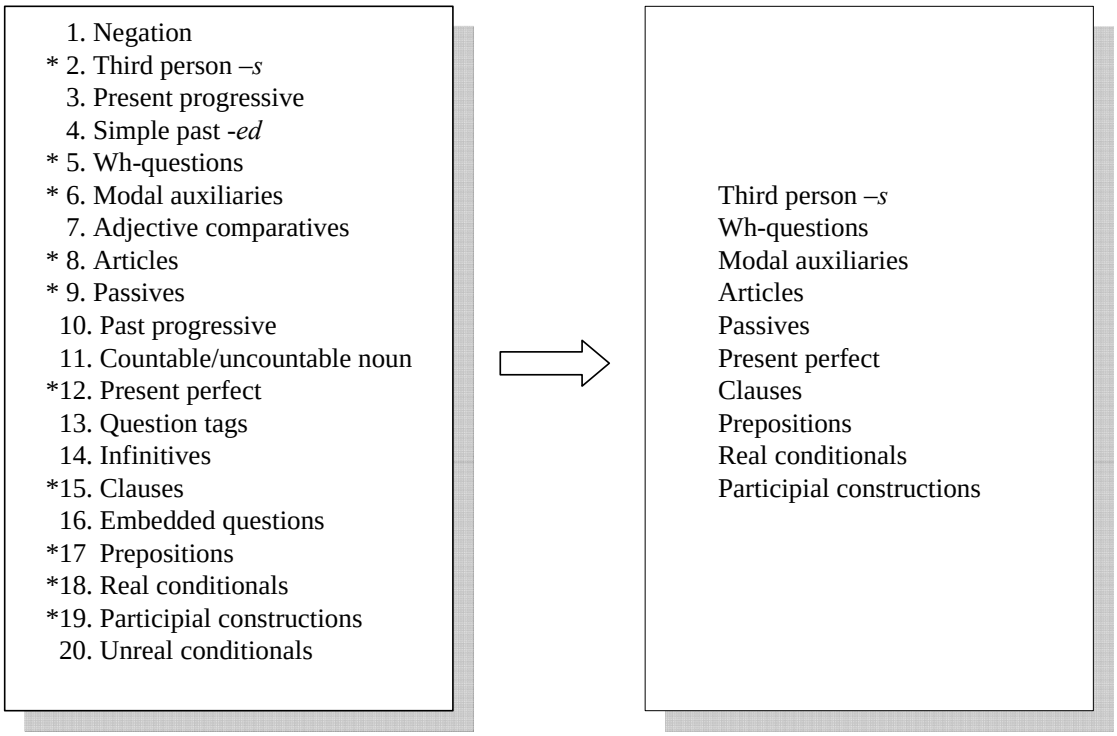


Figure 2.**Target features selected for the questionnaire and the grammatical ranking activity**.

For the ranking activity, the participants were given 10 cards for the ten structures. On each card, a metalinguistic term with its Chinese translation was provided, along with two sample sentences using the target structure; the metalanguage and sample sentences used for each individual structure were the same as those used in the questionnaire. Figure 3 illustrates the card format used in the ranking activity. In Figure 3, the Chinese translation of the metalinguistic term (i.e., present perfect) is given in parentheses.

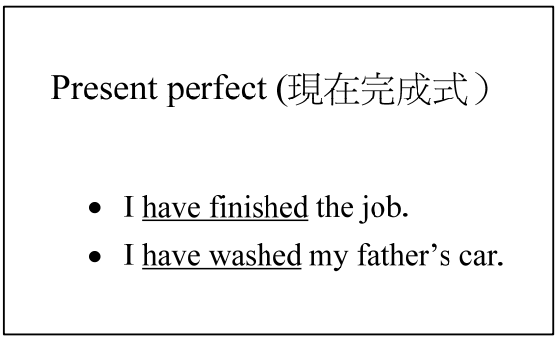


Figure 3. **The card format used in the grammatical difficulty ranking activity.**

The participants were asked to rank the target features from least to most difficult using the cards. After each participant completed the ranking, I asked him or her to explain the ranking using the following prompting questions.

- Why do you think that X is the least difficult?
- Why do you think that X is more difficult than Y?
- Why do you think that X is the most difficult?
- Anything else?

The prompted responses were all audio-recorded. The ranking results were also recorded for later quantitative data analysis.

Cloze activities.

To explore learners' perceptions of grammatical difficulty in more depth, I used two cloze activities (see Appendix F) and two stimulated recalls (see Appendix G). Cloze activities were used because they provide a context that helps learners to select the appropriate form of target verbs. The stimulated recall, conducted right after each cloze activity, allowed me to learn more about the decision-making process that lead to the participant's determination of the target features. The stimulated recall is discussed in more detail in the next section.

Two features of English grammar–the passive construction and the real conditional–were chosen for the cloze activities. One of the reasons for the choice of these two features was that they both suited a research need. As previously noted in Chapter Two, one purpose of the study was to investigate learners' perceptions of grammatical difficulty in relation to their L2

performance. In order to explore this, there was a need to identify two grammatical features that differ in terms of their level of difficulty. The passive construction and the real conditional were selected. A Wilcoxon test was conducted to compare the responses to the items representing these two features on the questionnaire. The results indicated a significant difference, $z = -11.70$, $p < .01$. $r = -.70$. The second reason for the choice of these two features was that they are both syntactic in nature, thus making these two features comparable in this respect (for more detailed discussion of these two features, see "Phase 3" subsection "Target Features*"*).

The cloze activities were adapted from existing cloze tests. One cloze activity uses a passage about the Statue of Liberty (Spada, et al., in preparation) with the target feature–passive construction.[17] There are six blanks in this cloze activity, each of which is immediately followed by a target verb in parentheses (for example, ____ (*give*)). The activity has six target verbs: *give*, *design*, *complete*, *ship*, *put*, and *visit.* Participants were asked to provide the appropriate form of the target verb in each blank. The other cloze activity uses a passage about a teacher's day at work (Valeo, 2010).[18] The target feature of this activity is the real conditional. The activity consists of filling in six blanks, four of which require the appropriate verb structure for the *if*-clause of the real conditional sentence, while the other two blanks require the appropriate verb structure for the main clause. The target verbs for the six blanks are *copula be*, *stay*, *tell,* and *get*. Given a concern about a task effect,[19] I counter-balanced the administration of the two cloze activities; that is, half of the participants were first provided with the cloze activity about the Statue of Liberty and the other half first received the cloze activity about a teacher's day at work.

The cloze activities were piloted with the first three students who participated in Phase 2 of the research. The purposes of piloting were to see (1) whether the activities would work to invoke a conscious decision-making process that could later be recalled, (2) whether content familiarity was a concern, and (3) how much time was needed to complete each activity. The pilot results showed that the students were able to recall the thoughts they had during the cloze activities, that content familiarity was not a problem, and that each cloze activity could be completed within 10 minutes. Given these results, the activities were used unchanged with the

[17] Material used and reprinted with the author's permission.
[18] Material used and reprinted with the author's permission.
[19] Each cloze test was immediately followed by a stimulated recall. It was assumed that the first simulated recall might influence the results of the second because the participants were more familiar with the procedure the second time around.

other participants involved in Phase 2 research. As the same cloze activities were used, it was decided to include the data from the pilot participants in the main study.

Stimulated recalls.

The stimulated recalls used in this study had two stages. Stage 1 aimed to explore what participants were thinking while doing the cloze activity. This stage was similar to the stimulated recalls discussed by Gass and Mackey (2000). Such stimulated recalls have the advantage of revealing what the learners have in mind during their language performance (Gass & Mackey, 2000). The first stage of the stimulated recall started with the prompting question, *Now you have finished the exercise, would you please tell me what you were thinking while you were doing the exercise?* After the participant's response, he or she was asked another prompting question: *What were you thinking when you were doing item X? You said XXX just now. Could you tell me more about it?* or *Anything else?*

Stage 2 took place right after Stage 1. Stage 2 aimed to explore what might contribute to the participants' decision-making while they were doing the cloze activity. The responses gathered at this stage were the ones used for the subsequent data analysis. Immediately after asking the questions to explore the participants' thinking, I then asked questions that expanded on the participants' responses obtained from the first stage of stimulated recall. Because the participants' responses varied, the questions asked in the second stage varied; however, they were all associated with learning the target features. The following excerpt illustrates what the second stage questions were like.

(This excerpt is extracted from the stimulated recall of one participant after she finished the cloze activity targeting the passive.)

In stage 1

Participant: I think I need to use the passive in this blank because I saw the word "by" in this sentence.

In Stage 2

Researcher: Just now you said that you thought you needed to use the passive because of the word "by". Why is that?

Both stimulated recall activities were conducted in Mandarin Chinese. All the stimulated recalls were audio-recorded with the participants' permission.

Data collection procedures.

Phase 2 of the research was conducted one week after the completion of Phase 1 research; it was conducted over the first two weeks of January 2009. During the one week interval between Phase 1 and Phase 2, I prepared the instruments and contacted the prospective participants. In Phase 2, at the students' convenience, I met with each of them individually on campus. Each of the meetings took less than an hour. During the meeting, after introducing myself, explaining Phase 2 of the research, and ensuring the consent form was signed, I asked the participant to do the activities in the following order:

1. A semi-structured interview (8 minutes or so)
2. A grammatical difficulty ranking activity (10 minutes or so)
3. A cloze activity (5 to 8 minutes)
4. Stimulated recall (5 to 8 minutes)
5. A cloze activity (5 to 8 minutes)
6. Stimulated recall (5 to 8 minutes)

All the activities were conducted in Mandarin Chinese, and with the participants' permission, they were all audio-recorded with a digital recording device.

Phase 3

Phase 3 involved the administration of the oral production tasks and the written metalinguistic task. I chose to explore learners' oral production and metalinguistic ability for two reasons. First, these two types of tasks are frequently used in L2 research. Thus, the findings of the current study regarding learners' performance on these two types of tasks can be compared with those discussed in the relevant literature. Second, having seen similar activities in their English classes, the participants in my study were familiar with the formats of these tasks. Thus, the influence of lack of test familiarity is of less concern.

In terms of the type of l2 knowledge measured by these tasks, the oral production tasks are argued to be conducive to greater retrieval of implicit knowledge for two main reasons. First, the nature of the one-on-one interaction with the researcher compels learners to spontaneously

produce language within time constraints. Second, the narrative nature of the task focuses learners' attention on meaning. In contrast, the metalinguistic task may create conditions favourable for greater use of explicit knowledge because the task focuses learners' attention on the language code itself. The oral production and written metalinguistic tasks will be described in more detail below. More discussion of the relationship between learners' task performance and the use of different types of L2 knowledge is provided in the discussion chapter.

Participants.

Oral production task participants.

The oral production tasks involved 27 students, including 19 who participated in Phase 2 of the research. When I first approached this student cohort to gauge their willingness to participate in Phase 3, everyone indicated interest; however, due to schedule conflicts, only 19 of them were able to participate in the oral production tasks. The other eight students were randomly selected from those who previously indicated interest in participating in Phase 2, but were not chosen. Of the 27 participants who completed the oral production tasks, 6 of them were male and 21 were female. In return for their time, each participant was given a gift coupon worth the equivalent of 3.3 Canadian dollars (100 New Taiwan dollars).

Written metalinguistic task participants.

The metalinguistic task initially involved 193 students from five of the classes that had participated in Phase 1. The selection of the classes was primarily based on the instructors' willingness to have their class participate in this phase of the study. Using the selection criteria indicated in the "Participants*"* section of Phase 1, 9 of the 193 students were excluded, thus leaving 184 students in the data analysis. Of these 184 participants, 41 were male and 143 were female. In terms of academic majors, 118 of them were from humanities, 46 were from social science, and 20 were from science.

Target Features.

The target features for the oral production tasks and the metalinguistic task were the passive construction and the real conditional in English.

Passive construction.

English has several types of passives such as the *be*-passive, the *get*-passive, and the *have*-passive. The current study focused on the *be*-passive, which is the prototypical passive in English. According to Celce-Murcia and Larsen-Freeman (1999), syntactically, the *be*-passive constitutes two parts: an auxiliary *be* and the past participle of a verb. The auxiliary *be* is determined by tense and aspect, whereas the past participle reflects the thematic verb (or the main verb). For example, in the sentence *The bridge was built in 1990*, the auxiliary *was* indicates past tense, and the past participle *built* reflects the thematic meaning of the verb *build*. In the same vein, in the sentence *Oil has been found in the Middle East, has been* indicates present perfect, and the past participle *found* represents the thematic meaning of the verb *find*. These two sample sentences also demonstrate that the passive permits the receiver of the action to take the subject position, and that the passive allows the absence of the agent of the action. However, if the agent of the action is present in a passive sentence, it is usually placed in the post-verbal position within a *by*-phrase. For instance, in the sentence *The cake was eaten by John*, the agent of the action, in this case, John, appears within a *by*-phrase that immediately follows the past participle of the verb *eat*. It should be noted that only transitive verbs are allowed in the passive. Given these syntactic attributes of the passive, accurate use of the *be*-passive entails, at least, knowing how to (1) differentiate the agent–receiver (or undergoer) of the action, (2) determine the tense and aspect, (3) construct the past participles of verbs, and (4) distinguish transitive verbs.

Citing Langacker (1987), Celce-Murcia and Larsen-Freeman (1999, p. 347) point out that the use of the passive is often semantically constrained. For example, passive sentences are more acceptable in the situations where

- the subject is more definite
  This song was sung by Amy Lin, but not *Songs were sung by Amy Lin.*
- with stative verbs, the object in the *by*-phrase is more indefinite
  Mary was liked by everybody, but not *Mary was liked by John*.
- "the verb denotes a physical action, as opposed to a state" (p. 347)
  The ball was kicked over the goalposts, but not *The ball was wanted by the other team* (Examples taken from Celce-Murcia & Larsen-Freeman, 1999, p. 347).

In addition, Celce-Murcia and Larsen-Freeman note that the passive is often used in situations where the agent of the action is unknown or unimportant; for example, in the sentences *Rice is grown in Taiwan* or *The supermarket was robbed last week*. The passive can also be used to draw attention to important information, increase coherence in writing, avoid overtly specifying the agent of the action, and avoid an overly long subject, to name a few. This suggests that learning the appropriate use of the passive entails not only semantic but also pragmatic knowledge. Accordingly, it is no surprise that the passive is problematic for L2 learners, even for advanced learners (Hinkel, 2002; S. Izumi & Lakshmanan, 1998; Larsen-Freeman, 1995).

Real conditional.

Conditional sentences are both syntactically and semantically complex. According to Celce-Murcia and Larsen-Freeman (1999), conditional sentences at the syntactic level contain a main clause and a subordinate clause. The subordinate clause is usually preceded by an adverbial subordinator *if* (thereafter, *if* clause). In most cases, the *if* clause can be positioned before or after the main clause; the variation in the position does not change the meaning of the sentence. For example,

If you leave, John will leave. *John will leave if you leave*.

Celce-Murcia and Larsen-Freeman also note that at the semantic level, conditional sentences can be broken down into various types. Table 1 reproduces the semantic hierarchy of conditional sentence types described by the authors, along with associated tense-modal sequencing and sample sentences (Celce-Murcia & Larsen-Freeman, 1999, pp. 548-552).

Table 1

Reproduction of the Semantic Hierarchy of Conditional Sentence Types

Conditional types	Sub-types	Semantic relationships	Tense-modal sequencing	Examples [a]

Conditional types	Sub-types	Semantic relationships	Tense-modal sequencing	Examples [a]
Factual conditionals	Generic factual conditionals	To express relationships that are based on physical law	*If* + simple present, simple present	• *If oil is added to water, it floats on top.* • *If you boil water, it vaporizes.*
	Habitual factual conditionals	To express relationships that are based on habit	*If* + simple present/past, simple present/past	• *If I wash the dishes, Sally dries them.* • *If Nancy said, "Jump!" Bob jumped.*
	Implicit inference conditionals	To express specific time-bound relationships	*If* + TAM_x[b], TAM_x	• *If smog can be licked in L.A., it can be licked anywhere.* • *If it is freezing outside, my roses are dying.*
	Explicit inference conditionals	To express specific time-bound relationships and are marked by a modal auxiliary for its inference	*If* + TAM, {*must/should*} VP	• *If someone is at the door, it must be Peter.* • *If anyone gets cake, it should be Rod.*
Future (predictive) conditionals	Strong condition and result	To express future plans	*If* + simple present, {*will/be going to*}VP	• *If it rains, I will stay home.* • *If Steve comes to class, he will get the answers to the quiz.*

Conditional types	Sub-types	Semantic relationships	Tense-modal sequencing	Examples [a]
	Degrees of weakened condition or result	To express future plans, but with less certainty	If {*should/happen to/should happen to*} VP, {should/may/might} VP	• *If you finish your vegetables, I might buy you an ice cream cone.* • *If it should rain, I may stay home.*
Imaginative (subjunctive) conditionals	Hypothetical conditionals	To express what the speaker perceives to be unlikely yet possible events or states in the *if* clause	*If* {*were* to V/ V + simple past}, *would* *If* {simple past/ present subjunctive}, *would*	• *If Joe had the time, he would go to Mexico.*
	Counterfactual conditionals	To express impossible events or states in the *if* clause	*If* {simple past/ present subjunctive}, *would* *If* + past perfect, *would have* V + *-en*	• *If my grandfather were alive today, he would experience a very different world.* • *If you had mowed my lawn, I would have paid you $5.*

Note: [a] Examples are used by the authors (1999, pp. 548-552). [b] TAM_x = Any possible combination of tense, aspect, modals, and phrasal modals.

As Table 1 shows, the conditional sentences are categorized into *factual conditional sentences*, *future (predictive) sentences*, and *imaginative (subjunctive) sentences*. The factual conditional sentences and the future (predictive) sentences are what I refer to as "real conditionals," while the imaginative (subjunctive) sentences are what I refer to as "unreal conditionals." I elected to use the terms "real conditionals" and "unreal conditionals" because

they appear more frequently in ESL/EFL textbooks and reference grammars, and, thus, would be familiar to the participants. As Celce-Murcia and Larsen-Freeman note, the tense-modal sequences listed above are by no means exhaustive. There is a large variety of tense-modal sequences that can be identified in spoken or written corpora (Hwang, 1979, cited in Celce-Murcia & Larsen-Freeman, 1999). However, Hwang's corpus study reports that the two most frequent types of conditional sentences are "generic factual conditionals" ("*If + simple present, simple present*") and "future predictive conditionals" ("*If + simple present, {will/be going to} VP*"). These two types of conditional sentences account for approximately 30% of all the types of conditional sentences that Hwang identifies in her corpora. Thus, according to Hwang's findings, real conditionals are more frequently used than unreal conditionals.

In the present study, only real conditionals were used in the tests. The tense-modal sequences targeted were the following:

If + simple present, (then) simple present →*If John <u>is</u> on time, he <u>walks</u> to work.*

If + simple past, (then) simple past → *If I <u>had</u> time, I <u>walked</u> home*.

If + simple present, (then) future → *If it is not too late, John <u>will walk</u> to work*.

If + simple present, (then) modal → *If John <u>leaves</u> on time, he <u>may walk</u> to work.*

If + simple present, (then) imperative → *If you have time, <u>walk</u> with me*.

Research Instruments.

Oral production tasks.

Three picture-cued oral production tasks were designed to elicit communicative use of the two target features. Each picture-cued task was designed to elicit the use of one target feature.

Two of the three tasks targeted the passive construction in English. I created one of the tasks targeting the passive construction (see Appendix H). This task, entitled "What Might the Book Say," consists of eight pictures. The first is an introductory picture; the second is used to ensure that the students understand the vocabulary used in the following six pictures. Each of the following six pictures is about some historical fact; a verb and information about the historical fact depicted are provided in each picture. Figure 4 presents two examples of these pictures. This task requires that students use the provided verb to describe what the book might say about each

picture. The target verbs used are *send*, *invent*, *build*, *arrive*, *find*, and *complete*. The verb *arrive* functions as a distracter. For each picture, the students were prompted with the question "What might the book say about this picture?"

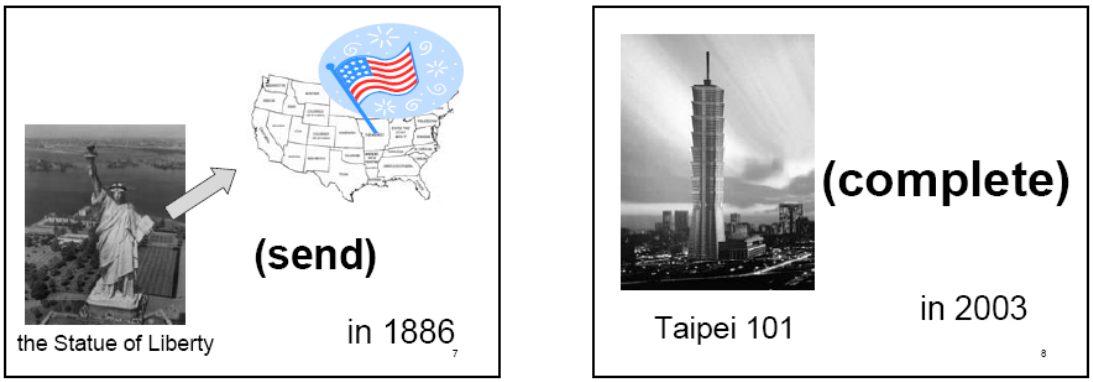


***Figure 4*. Examples of the pictures in the "Book" task.**

The other task targeting the passive construction (see Appendix I) was adapted from an oral production task about a package lost in the mail (Spada, et al., in preparation).[20] This task consists of eight pictures plus one introductory picture. In each of the eight pictures, a verb is provided. Students are required to describe each picture using the provided verb. The target verbs used in the pictures are *mail*, *send*, *return*, *put*, *arrive*, *deliver*, and *call*. The verbs "arrive" and "call" are used as distracters. Figure 5 contains two pictures from this task; for each picture, I prompted the students by saying, "What happened to the package?"

[20] Materials used and reprinted with the author's permission.

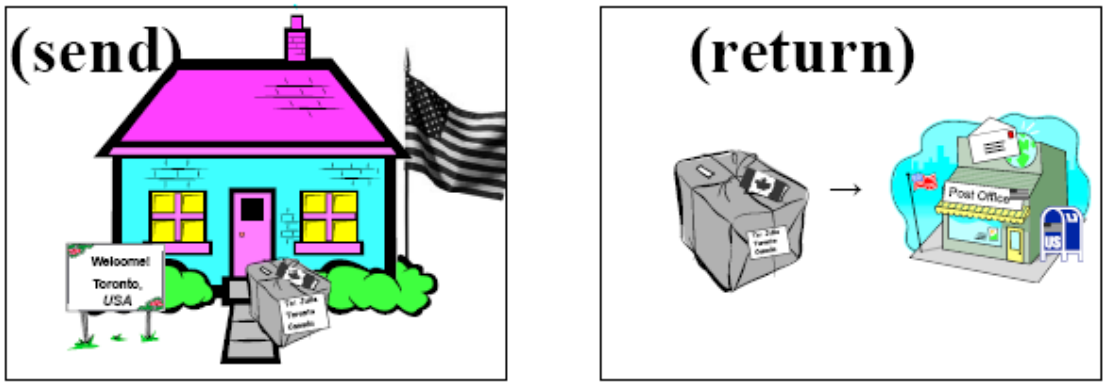


***Figure 5*. Examples of the pictures in the "Package" task.**

I designed the task targeting real conditionals (see Appendix J). The task consists of 9 pictures. The first 4 pictures introduce information about special offers provided by a travel agency. The other 5 pictures ask the students to pretend to be travel agents, and to give the information provided in each picture. Figure 6 illustrates two examples of these five pictures. Each of them provides one piece of information about a special offer. Students are expected to give the information by starting their sentence with "If you. " The prompting sentence used for each picture is, "What would you say here?"

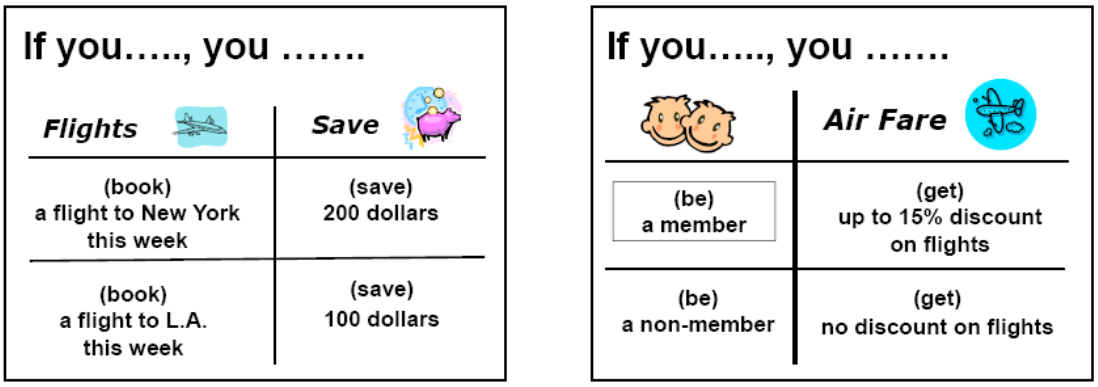


***Figure 6*. Examples of the pictures in the real conditional task.**

The three oral picture-cued tasks were piloted to ensure that they could successfully elicit the use of the target features.[21] The pilot participants included three native speakers of English and three students who were comparable to the participants in the main study. Changes were made where necessary. The final pilot results indicated that the tasks could successfully elicit the use of the target features, and the time needed for completing all three tasks is less than 10 minutes.

Written metalinguistic task.

The written metalinguistic task consists of three sections: error correction, error identification and error explanation. While it could be convincingly argued that explanation is more metalinguistic in nature than error identification and correction, for the purpose of simplicity, I have chosen this label for this written test (see Appendix K). The task consists of 45 items: 15 items target the passive construction, another 15 items target the real conditional, and the other 15 items are distracters. Each of the 45 items contains a sentence with a grammatical error. The participants were required (1) to identify the ungrammatical element in each sentence, (2) to provide the correct form, and (3) to explain their correction. The participants were allowed to use English or Chinese, or both, in their explanations. Metalinguistic terminology was not required. The following are two examples of the test items.

1. They drive to Taipei twice last week.

The ungrammatical part is

_____

The correct form is

_____

It is ungrammatical because

_____

_____

[21] The oral production task by Spada, et al. (in preparation) has high reliability, with Cronbach's alpha index of .86.

2. Many messages were receiving yesterday.

The ungrammatical part is

_____

The correct form is

_____

It is ungrammatical because

_____

_____

The target features for the 15 distracter items are third person *–s*, wh-questions, and adjective comparatives. These distracter items were not included in the subsequent data analysis. Eight of the 15 items targeting the passive construction were taken from the error correction test created by Spada et al. (in preparation).[22] I created the other 7 items. Of these 15 items, 5 items contain errors of missing *copular be* (for example, *Many messages <u>received</u> yesterday*), while the other 10 items contain errors in past participles (for example, *Rules are <u>making</u> by the school*, *The test will be <u>give</u> orally*). With regard to tense and aspect, 8 of the passive items involved a simple tense-aspect, 4 of them involved a complex tense-aspect, and three of them involved the use of a modal auxiliary (see Table 2). Here, simple tense-aspect refers to simple present or past tense, whereas complex tense-aspect refers to present perfect.

[22] Materials used and reprinted with the author's permission.

Table 2

Items Targeting the Passive in the Written Metalinguistic Task

	Item No.	Item
Item involving simple tense-aspect	8	Rules are making by the school.
	29	Vitamins were discovering in the early 1900's.
Item involving complex tense-aspect	22	Many people have been save since insulin was discovered.
	26	His heart has been breaking three times this year.
Item involving the use of modal auxiliary	5	The test will be give orally.
	20	The contract must be sign by the boss.

Among the 15 items targeting the real conditional, 8 items were taken from the error correction test by Valeo (2010),[23] and I created the other 7 items. Of the 15 items, 10 have an *if* clause in initial position, while the other 7 have a main clause in initial position. Of the 10 items that start with an *if* clause, 5 contain an error in the verb of the *if* clause (See Example A below), while the other 5 contain an error in the verb of the main clause (see Example B below).

Example A

If it will snow tomorrow, they will cancel the meeting.

Example B

If children eat breakfast, they having more energy.

[23] Materials used and reprinted with the author's permission.

With regard to the 5 items starting with the main clause, 2 of them contain an error in the verb of the main clause, while the other 3 contain an error in the verb of the *if* clause. Example C illustrates the former items, while Example D illustrates the latter.

Example C

You must taking a taxi if you miss the bus.

Example D

My family will go to the zoo if the weather will be nice tomorrow.

The metalinguistic task was piloted in one class of students (n = 50) who were comparable to the participants in the main study. The students were told to take as much time as they needed to do the test. The pilot results showed that, on average, the task took an hour to complete, with the time ranging from 50 minutes to 70 minutes. Overall the reliability of the task was good, with a Cronbach's alpha index of .86 and .83 respectively for the items targeting the passive construction and the items targeting the real conditional. Individual items that had comparatively lower reliability levels were revised. The revised version of the task was then piloted with learners in Class 1 (n = 30), which was the class that participated in the piloting of the questionnaire and the proficiency test. The reliability of the revised version was .80 and .86 respectively for the items targeting the passive construction and the real conditional. Given the high level of reliability, no further revisions were made and the task was administered to the other four classes participating in Phase 3 of the study.

Data collection procedures.

Phase 3 of the research was conducted from the fourth week of February to the second week of March 2009. It was conducted six weeks after the completion of Phase 2. Because this six-week interval included a one-month winter break, the participants did not receive much additional EFL instruction between Phase 2 and Phase 3. During this six-week interval, I prepared and piloted the oral production tasks and the metalinguistic task, and I contacted the interested students and teachers to arrange to administer the tasks.

During the fourth week of February and the first week of March 2009, the oral production tasks were administered on campus at the participants' convenience on a one-on-one basis. I met

with each participant individually for 10 to 15 minutes to conduct the tasks. The oral production tasks started after the signed consent form was collected. The three oral production tasks were administered in a counter-balanced order to avoid a potential task order effect. The oral production tasks were conducted in English and, with the students' permission, was audio-recorded with a digital recording device.

During the second week of March 2009, the written metalinguistic task was administered to four of the classes that completed the questionnaire and the proficiency test in Phase 1. The task was administered under my supervision during regular class time. Before I administered the task, I collected the consent forms from the students. The students were told to take as much time as they needed to complete the task. The tests were collected only when all the students had finished. On average, each class took approximately an hour to finish the task, with the time ranging from 50 to 75 minutes.

Summary

The study consisted of three distinct phases in which multiple methods were employed to explore different Research Questions, and different numbers of participants were involved. Figure 7 summarizes the data collection procedures.

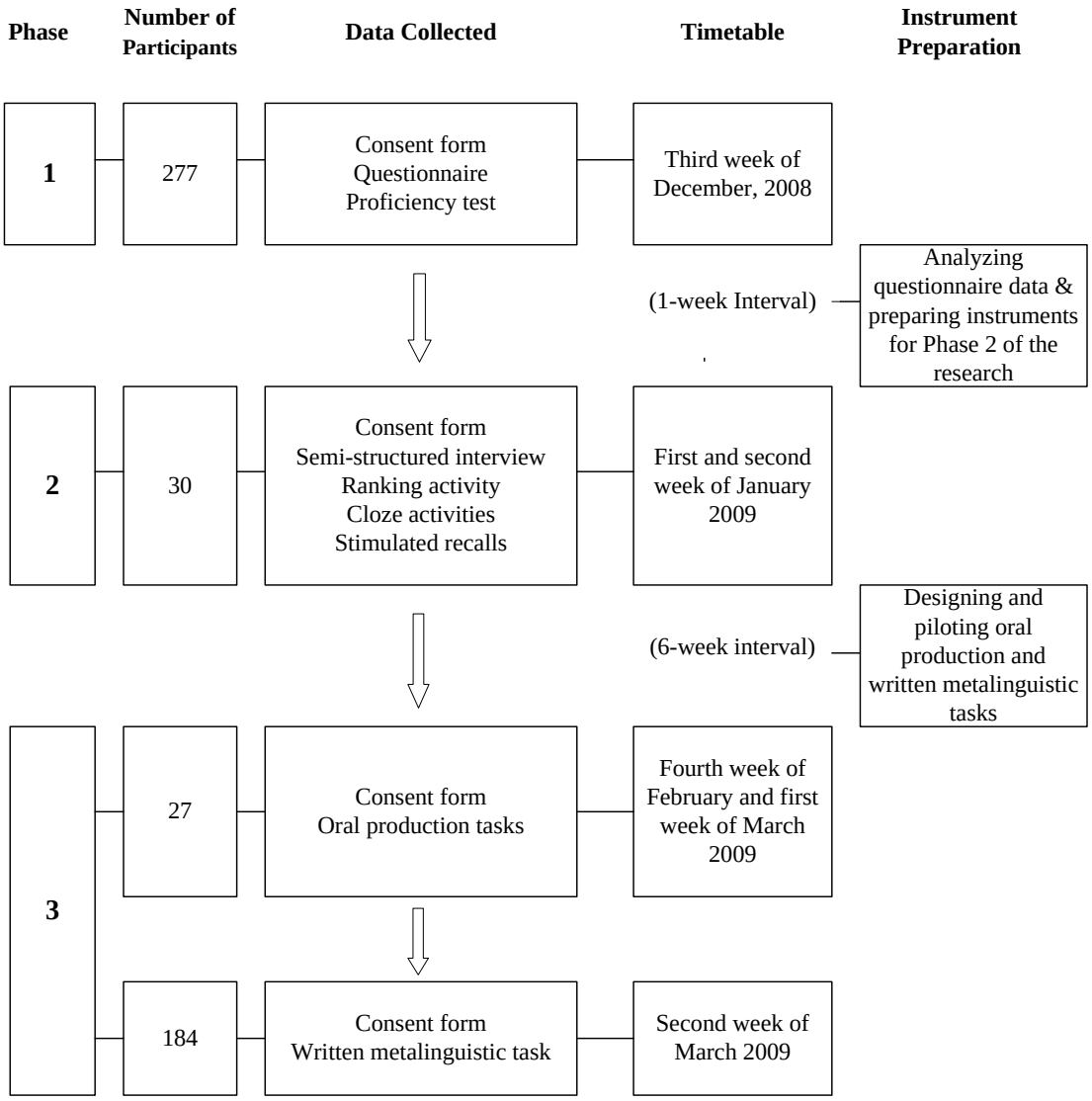


Figure 7. **Summary of data collection.**

Chapter Four
Data Analysis and Results for Research Questions One and Two

This chapter presents the analyses and results related to Research Questions 1 and 2. It begins with a discussion of the validity and reliability testing of the quantitative and qualitative data. This is followed by a presentation of the procedures used in the data analysis and the findings investigating the following two Research Questions:

1. Which features of English grammar do university-level Chinese EFL learners perceive as easier, and which as more difficult, to learn?
2. What influences Chinese EFL learners' perceptions of grammatical difficulty?

Validity and Reliability of the Quantitative and Qualitative Data

Validity.

Quantitative data were primarily gathered from the administration of the questionnaire. To ensure its validity, two steps were taken. First, expert opinion was sought to enhance content validity of the questionnaire. This was obtained through consultations with my thesis committee members and my colleagues. Based on their comments and suggestions, the questionnaire was revised several times before it was piloted. Second, seeing validation as an *ongoing* process of seeking evidence for the validity of a particular interpretation of test results (Bachman, 2004; Messick, 1989), validation measures were taken not only on the instrument design, but also on the data analysis and data interpretation. The latter was achieved by (1) cross-validating the questionnaire findings with the ranking activity results, and (2) where appropriate, discussing the questionnaire findings with reference to the relevant qualitative findings.

Qualitative data were gathered through the administration of a semi-structured interview, a grammatical ranking activity, two cloze-activity-based stimulated recalls, and the reflective items of the questionnaire. For qualitative research, the concept of validity can be broken down into what Lincoln and Guba (1985) term *credibility* and *transferability*. Credibility is the extent to which the results of a qualitative study are believable, and transferability is the extent to which the results can be generalized to other, similar contexts. To enhance the credibility and

transferability of the qualitative data, I have provided contextualized descriptions of the examined issues, and used several types of triangulation: *data triangulation*, *investigator triangulation*, and *methodological triangulation* (Brown & Rodgers, 2002). Data triangulation was achieved by including multiple participants (in this case, 30 participants). Investigator triangulation was accomplished by asking a fellow doctoral candidate to code the qualitative data and these were compared with my coding. Methodological triangulation was achieved by using multiple methods to collect the qualitative data (i.e., through the ranking activity, the cloze-activity-based stimulated recalls, and the "reflective items" on the questionnaire).

Reliability.

The internal-consistency reliability of the questionnaire was assessed by computing Cronbach's alpha. The results showed that, overall, the questionnaire has a high degree of reliability, Cronbach's alpha .93.

To ensure intra-coder reliability of the qualitative data, I coded all the data twice. There was a two-week interval between the two rounds of coding. Any mismatches found between the two rounds of coding were double checked. Where necessary, the themes were revised and were specified more clearly. The intra-coder reliability was computed using a subset of the data, k = .94. For the inter-coder reliability, a subset of the data was coded by a fellow doctoral student, using the finalized five themes (see *Results for Research Question 2*). The inter-coder reliability is good, k = .89. Differences in coding were reconciled through discussion.

Data Analysis for Research Question 1

The quantitative data from the questionnaire and the ranking activity were analyzed to address Research Question 1: Which features of English grammar do university-level Chinese EFL learners perceive as easier, and which as more difficult, to learn? I used the SPSS for Windows 15.0 to analyze all the quantitative data. I started by computing the descriptive statistics of the students' responses to all items in the questionnaire in order to obtain an overview of the data. Using SPSS Frequency, I conducted a check for the accuracy of the data entry, missing data, skewness, and kurtosis for the questionnaire dataset. I also inspected the minimum and maximum values, means, and standard deviations of each of the items for

plausibility. Following the inspection, I assessed the internal consistency of the 20 closed-ended items by computing Cronbach's alpha coefficients.

To examine students' perceptions of the grammatical difficulty of the 20 features represented by the closed-ended items, the items were ranked in ascending order based on the value of their mean scores. With a view to validating the ranking results, three tests were performed: a Spearman correlation test, a principal components analysis (PCA), and a Pearson product-moment correlation test. Spearman correlation coefficients were computed on the 20 items to inspect the interrelationships of the students' responses to the items. A non-parametric correlation test was used in light of the skewed distributions of the data to be examined. Using the Bonferroni approach to control for Type I errors across the 20 correlations, a p value of less than .0025 (.05/20 = .0025) was required for significance. To further explore the interrelationships, an exploratory PCA was performed on the Spearman correlation matrix. The results of Promax rotations were examined. A Pearson product-moment correlation test was used to compare the questionnaire results with the ranking activity results. Three sets of mean scores were used for this comparison: the descriptive statistics of (1) the questionnaires completed by the 277 questionnaire respondents, (2) the questionnaires completed by the 30 participants who did the ranking activity, and (3) the quantitative results generated by the ranking activity. Because the ranking activity involved only 10 features, only the mean scores for those 10 features in the questionnaire were examined. The ranking activity required each student rank the 10 features from least to most difficult, resulting in a set of ranking orders from 1 to 10, with 1 indicating the least difficult and 10 indicating the most difficult. The ranking activity involved 30 students, thus producing 30 sets of ranking orders. The mean scores were computed on these 30 sets.

Upon completion of the validation tests, I examined and compared the items in the lower ranking with those in the higher ranking to explore what might explain the grammatical difficulty ranking.

Results for Research Question 1

Table 3 reports the descriptive statistics, item-remainder correlations, and Cronbach's alpha for each item. The 20 items are ranked by their mean scores (from lowest to highest).

Table 3

Descriptive Statistics of the Questionnaire

Items	*n*	*M*	*SD*	Mode	Skew [a]	Reliability	
						Corrected item total correlation	Cronbach's alpha if item deleted
Negation	277	1.17	.57	1	4.37	.54	.93
Third Person -*s*	277	1.24	.57	1	3.00	.55	.93
Present Progressive	276	1.29	.65	1	2.95	.65	.93
Simple Past -*ed*	277	1.32	.71	1	2.92	.64	.93
Wh-questions	277	1.47	.80	1	2.18	.64	.92
Modal Auxiliaries	277	1.50	.88	1	1.99	.59	.93
Adjective Comparatives	276	1.76	.83	1	.98	.65	.92
Articles	276	1.82	1.09	1	1.56	.54	.93
Passives	277	1.84	.96	1	1.07	.73	.92
Past Progressive	277	1.88	1.10	1	1.24	.60	.93
(Un)countable Noun	277	2.00	1.11	1	1.07	.55	.93

Items	*n*	*M*	*SD*	Mode	Skew [a]	Reliability	
						Corrected item total correlation	Cronbach's alpha if item deleted
Present Perfect	277	2.00	1.08	1	1.14	.58	.93
Question Tags	277	2.04	1.05	2	1.08	.64	.92
Infinitives	276	2.11	1.08	2	.96	.71	.92
Clauses	277	2.35	1.20	2	.78	.66	.93
Embedded Questions	277	2.51	1.14	2	.50	.64	.92
Prepositions	277	2.87	1.40	2	.43	.62	.93
Real Conditionals	277	3.06	1.34	3	.31	.62	.93
Participial Construction	277	3.11	1.40	3	.30	.66	.92
Unreal Conditionals	277	3.32	1.28	3	.13	.61	.93

Note: [a]. Standard error of skewness is .15.

Looking at the mean scores, we see that 17 out of the 20 items have a mean score below an averaged mean score of 3 (on a continuum from 1 to 6), while three remaining items have a mean score of around 3. The lowest mean score of the 20 items is 1.17, and the highest is 3.32. A reading of standard deviations indicates that the range of the standard deviations is small (from .57 to 1.14) and that less difficult features (i.e., features with lower mean scores) tend to

have a smaller standard deviation than more difficult features (i.e., features with higher mean scores). Positive skewed distributions are observed on students' responses to all but the item "unreal conditionals."[24] The internal consistency of each individual item was evaluated by inspecting the values of *alpha if item deleted* and *item-remainder correlations*, which together show that all the items have high internal-consistency reliability.

A Spearman correlation test was performed to examine the interrelationships of students' responses to the 20 items. The results (see Appendix L) show that 19 out of the 20 correlations are statistically significant. The only exception is the correlation of negation with unreal conditionals, $\rho\,(275) = .176$, $p = .003$. Scanning through all the correlations of these 20 items, it was revealed that the correlation coefficients range from .18 to .71, and that more difficult features tend to have higher correlations with more difficult features, but lower correlations with less difficult features, and vice versa. For example, two features, real conditionals and unreal conditionals, ranked as the most difficult features, have high correlations with each other, $\rho\,(275) = .71$, $p < .001$. However, they have low correlations with negation and third person *–s*, the two features ranked as the least difficult (see Table 4). These results support the trustworthiness of the quantitative data obtained from the questionnaire.

Table 4

Spearman Correlation Coefficients of the Four Structures

	Negation	Third person *–s*
Real conditionals	.21(***)	.23(***)
Unreal conditionals	.18	.22(***)

Note: *** $p < .001$, 2-tailed.

An exploratory PCA was conducted to further explore the interrelationships of the responses to the 20 items. The suitability of the data for this analysis is supported by the observations that (1) many of the correlation coefficients were above .32, (2) the Kaiser-Meyer-

[24] Responses to the item of "unreal conditionals" have a normal distribution.

Oklin value (.93) exceeded the recommended value of .6, and (3) Bartlett's Test of Sphericity reached statistical significance (Field, 2005). I first performed PCA on the Spearman correlation matrix to estimate the number of components. I then used three criteria, suggested by Tabachnick and Fidell (2007), to determine the number of the components to rotate: Kaiser's recommendation of retaining components with eigenvalues greater than 1, the scree plot of eigenvalues, and the interpretability of the component solution. Three components had eigenvalues greater than 1 and explained 58.34% of the item variance. The scree plot showed a sharp descent after the first component, a small drop after the second component, and an even smaller one after the third component (see Figure 8). Therefore, I considered two- and three-component solutions for the subsequent analysis.

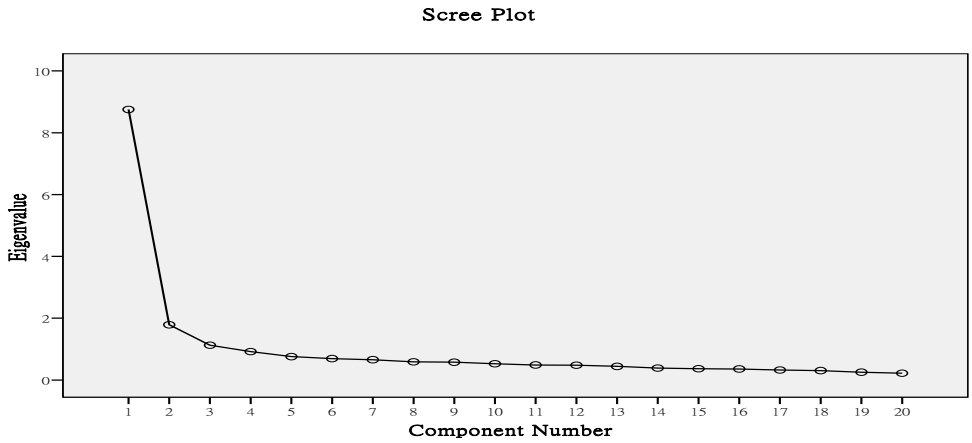


***Figure 8*. Scree plot of eigenvalues for each component.**

I used the Principal Component extraction technique with Promax rotation for the two- and three-component solutions. Following Field (2005), an absolute value of .40 was used; all factor loadings with the value of less than .40 were suppressed. The three-component solution resulted in only a few items loading higher than .40 on the third component. The two-component solution indicated that the majority of the items loading higher than .40 are more or less equally distributed on either the first or the second component. Considering the interpretability of the component solution, therefore, a decision was made to use the two-component solution. The two

components in the final solution account for 52.70% of the variance. After rotation, Component 1 explains 43.76% of the item variance (eigenvalue = 8.75), while Component 2 contributes 8.94% (eigenvalue = 1.79). Table 5 presents the component loadings.

Table 5

Component Loadings Using Principal Factors Extraction and Promax Rotation

	Component	
	1	2
Unreal Conditionals	**.95**	
Real Conditionals	**.93**	
Participial Construction	**.84**	
Embedded Questions	**.74**	
Clauses	**.70**	
Infinitives	**.62**	
Question Tags	**.51**	
Passives	**.47**	
Prepositions	**.44**	
Adjective Comparatives	.39	.37
Negation		**.87**
Simple Past		**.84**
Third Person *-s*		**.83**
Present Progressive		**.74**
Modal Auxiliaries		**.68**
Articles		**.62**
(Un)Countable Nouns		**.52**
Wh-questions		**.49**
Past Progressive		**.45**
Present Perfect	.35	.36

In PCA, loadings are referenced to interpret a component. As a rule of thumb, the greater the loading generated from using oblique rotations, the better interpretation it will yield. Comrey and Lee (1992, cited in Tabachnick & Fidell, 2007) suggest that loadings exceeding .71 are considered excellent, .63 are very good, .55 good, .45 fair, and .32 poor. To be conservative in interpreting the variables, I examined only the features with loadings greater than. 63. I compared the features with loadings greater than .63 in Component 1 and in Component 2. It was found that the features with loadings greater than .63 in Component 1 are those ranked as the most difficult features (i.e. the features with highest mean scores), while the features with loadings greater .63 in Component 2 are those ranked as the least difficult features (i.e., the features with lowest mean scores). I interpreted these findings as further support to the trustworthiness of the ranking results from the questionnaire.

The questionnaire ranking results were also cross-validated, using a Pearson product-moment correlation test, with the results of the ranking activity. As noted earlier, three sets of mean scores were used for the comparison: the descriptive statistics of (1) the questionnaires completed by the 277 questionnaire respondents, (2) the questionnaires completed by the 30 participants who did the ranking activity, and (3) the quantitative results generated by the ranking activity. Table 6 presents the three sets of mean scores that were used for the comparison. In Table 6, “Questionnaire Dataset A” refers to the mean scores of the questionnaire completed by all the questionnaire respondents, while “Questionnaire Dataset B” refers to the mean scores of the questionnaire completed by the 30 respondents who did the ranking activity. The 10 mean scores in these two datasets were obtained by computing students’ responses to the closed-ended items in question. The mean scores of the ranking activity were obtained by computing the 30 sets of the ranking order gathered from the ranking activity.

Table 6

Comparison of the Three Sets of Mean Scores

	Questionnaire Dataset A (n = 277)		Questionnaire Dataset B (n = 30)		Ranking activity (n =30)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Third person -s	1.24	.57	1.27	.58	1.53	.73
Wh-questions	1.47	.80	1.50	.68	2.70	1.26
Modal auxiliaries	1.50	.88	1.43	.73	3.80	1.56
Articles	1.82	1.09	2.07	1.20	3.17	1.98
Passives	1.84	.96	1.93	.87	5.83	1.80
Present perfect	2.0	1.08	2.17	1.09	7.57	1.85
Clauses	2.35	1.20	2.67	1.32	6.93	1.62
Prepositions	2.87	1.4	3.27	1.44	6.83	1.90
Real conditionals	3.06	1.34	3.37	1.61	8.13	1.63
Participial construction	3.11	1.40	3.37	1.38	8.57	1.79

Table 7 presents the results of the Pearson correlation test that was conducted to explore the correlations among the three sets of mean scores. The results show that the three sets of mean scores were all highly correlated, r = 1.00, p = .000, r = .88, p = .001, r = .86, p = .001,

respectively. These high correlations offer another piece of evidence in support of the trustworthiness of the ranking results from the questionnaire.

Table 7

Correlations of the Three Sets of Mean Scores

	Questionnaire ranking results (n = 277)	Questionnaire ranking results (n = 30)	Ranking activity results (n = 30)
Questionnaire ranking results (n = 277)	–		
Questionnaire ranking results (n = 30)	1.00(**)	–	
Ranking activity results (n = 30)	.88 (**)	.86(**)	–

Note: ** p < .01, 2-tailed.

The results of the three tests–the Spearman correlation test, the PCA test, and the Pearson correlation test–all indicate the trustworthiness of the ranking results. Given this finding, I examined and compared the features lower in the ranking with those higher in the ranking. Because the descriptive statistics shows that the mean score differences for some features are small, I compared only the six features with the lowest mean scores (see the left side of Table 8) with the six features with the highest mean scores (see the right side of Table 8) in order to look for possible trends.

Table 8

Comparison of the Features with the Lowest Mean Scores and Those with the Highest Mean Scores

Features with the lowest mean scores		Features with the highest mean scores	
Feature	***M***	**Feature**	***M***
Negation	1.17	Clauses	2.35
Third Person -*s*	1.24	Embedded Questions	2.51
Present Progressive	1.29	Prepositions	2.87
Simple Past -*ed*	1.32	Real Conditionals	3.06
Wh-questions	1.47	Participial Construction	3.11
Modal auxiliaries	1.5	Unreal Conditionals	3.32

Data Analysis for Research Question 2

Qualitative data were analyzed to address Research Question 2 – What influences Chinese EFL learners' perceptions of grammatical difficulty? As noted earlier, the qualitative data were gathered from the administration of a semi-structured interview, a grammatical ranking activity, two cloze-activity-based stimulated recalls,[25] and the responses to the reflective items of the questionnaire. I first transcribed all the qualitative data and then read through the transcribed data to get an overall sense of them. After that, I analyzed the data using NVivo 8 (QSR International, 2009). NVivo 8 is a qualitative data analysis computer program, which has

[25] To address Research Question 2, only responses gathered at stage 2 of the stimulated recalls were analyzed.

the advantage that it allows for easy storing, manipulating, searching, and retrieving of transcribed data.

Using NVivo 8, I located "units of information" from the dataset for data coding (Lincoln & Guba, 1985). For each unit of information, I used a chunk of the transcript that was self-contained and expressed a unique, informative message. The units of information were selected using the principles suggested by Lincoln and Guba (1985):

> First, it should be heuristic, that is, aimed at some understanding or some action that the inquirer needs to have or to take . . . . Second, it must be the smallest piece of information about something that can stand by itself, that is, it must be interpretable in the absence of any additional information other than a broad understanding of the context in which the inquiry is carried out. Such a unit may be a simple factual sentence . . . . It may be as much as a paragraph. (p. 345)

For the analysis of the semi-structured interview data, I simply summarized students' responses to the questions asked in the interviews about their prior grammar learning experience, enjoyment of grammar learning, and attitudes toward the usefulness of grammar learning. However, for the analysis of the data from the ranking activity, the stimulated recalls, and the questionnaire, I searched for emerging themes through inspections of the coded data. The whole set of data was analyzed recursively until no new themes could be found. Then the identified themes were explored for their interconnections. Based on the interconnections found, the themes were further modified.

Results for Research Question 2

Background profiles of the 30 participants.

Before reporting on the results for Research Question 2, the background of the 30 students from whom the primary qualitative data were collected (see Table 9) is described. The names used are all pseudonyms. The background information was primarily collected from the students' responses in the questionnaire and in the semi-structured interviews. The background information reported in Table 9 consists of information about the students' sex, length of time studying English, self-assessment of grammar knowledge, and attitudes toward the usefulness of grammar learning. The last piece of information was collected from the semi-structured interviews, and the others were collected from the questionnaire.

Table 9

Background Profiles of the 30 Participants

Name	Sex	Years of English learning	Self-assessment of grammar knowledge	Attitude toward the usefulness of grammar learning
Abby	Female	7	Good	Useful for writing as well as reading
Beth	Female	6	Good	Useful for writing and reading
Charlie	Male	7	Good	Useful for writing as well as reading
David	Male	7	Not very good	Useful for writing and reading
Elvira	Female	8	Good	Useful for writing
Fiona	Female	9	Good	Useful for writing and reading
Goldie	Female	8	?	Useful for writing as well as reading
Hank	Male	9	Not very good	Useful for writing and reading
Howard	Male	9	Very Good	Useful for all four skills
Jack	Male	9	Not very good	Not very useful
Jacky	Male	8	Good	Useful for writing as well as reading
Jerry	Male	11	Not very good	Useful for writing

Name	Sex	Years of English learning	Self-assessment of grammar knowledge	Attitude toward the usefulness of grammar learning
John	Male	10	Not very good	Useful for reading
Joyce	Female	12	?	(Not clear)
Judy	Female	?	Not very good	Useful for writing and tests
Kay	Female	11	Not very good	Useful for reading
Kelly	Female	10	Very Good	Not very useful
Lily	Female	6	Good	Useful for writing
Lori	Female	7	Good	Useful for reading and writing
Mai	Female	6	Good	Useful for reading as well as writing
Mille	Female	6	Not very good	Useful for reading
Nicole	Female	7	Not very good	Useful for writing
Paula	Female	7	Not very good	Useful for writing and reading
Peter	Male	11	Good	Useful for writing and reading
Sandy	Female	11	Not very good	Useful for writing as well as reading

Name	Sex	Years of English learning	Self-assessment of grammar knowledge	Attitude toward the usefulness of grammar learning
Stella	Female	8	Good	Useful for writing and reading
Sunny	Female	12	Good	Useful for all four skills
Thomas	Male	10	Not very good	Useful for writing, reading as well as speaking
Tim	Male	10	Not very good	Not very useful
Tin	Female	7	Good	Useful for writing

Note: "?" indicates "not reported."

As Table 9 shows, 19 of the participants were female and 11 were male. The length of time spent learning English ranged from 6 to 12 years. The majority of the students reported that they did not have much exposure to English outside of the classroom. Thirteen students self-assessed their grammar knowledge as "not very good." Another 13 students assessed their grammar knowledge as "good," and two students assessed their grammar knowledge as "very good." With regard to students' attitudes toward the usefulness of learning grammar, three students (Kelly, Tim, and Jack) did not think that learning grammar was useful to English learning, while the other students acknowledged the usefulness of grammar in learning English, especially for English writing and reading.

Themes related to learners' perceptions of grammatical difficulty.

Five themes pertinent to learners' perceptions of grammatical difficulty were identified through the data analysis: *knowledge of syntactic constituents, knowledge of semantics, knowledge of pragmatics, previous grammar teaching and learning,* and *L1 influence*. These are discussed in the following.

Knowledge of syntactic constituents.

"Knowledge of syntactic constituents" refers to knowledge of morphosyntactic or syntactic aspects of a grammar feature. Students' responses such as "*the rules of thumb for third person -s are quite simple"*[26] or "*the passive is composed of only auxiliary be and the past participle*" were coded as "knowledge of syntactic constituents." A feature with simple grammatical rule of thumb was considered to be less difficult to learn than a feature with complex rules. Examples of the former are students' perceptions of the learning difficulty of third person *–s* and modal auxiliaries. The majority of the 30 students commented that third person *–s* is easy to learn because formulation of this feature requires a simple rule of thumb. Likewise, modal auxiliaries were perceived to be an easy feature for the same reason. One participant, Hank, remarked that modal auxiliaries are easy to learn because, "this feature has only one simple grammatical rule of thumb–that is, using the base form of a verb after a modal auxiliary." Participants Kay, Elvira, Howard, Mille, Lily, Sunny, and David all perceived modal auxiliaries as easy features for the same reason. Conversely, many of the 30 participants considered the participial construction to be more difficult to learn due to its complex grammatical rules of thumb. Jerry's comment about the difficulty of this feature is typical of the responses of the participants:

> This feature is difficult for me to learn because one needs to know rules for using the active voice or passive voice in order to decide whether to use *–ing* or *p.p.* [past participle]. One also needs to know the rules determining when to omit the subject in the sentence using participial construction. Too many rules! Too many rules need to be memorized in order to arrive at the correct form of participial construction.

However, although some of the target features were perceived to be more difficult to learn than others in terms of grammar rules, many of the students also commented that overall, they perceived the 10 target features selected for the ranking activity to be relatively easy to learn. Notwithstanding, some students also pointed out that knowledge of the features that is not represented by the illustrative sentences was difficult for them. To illustrate, despite the fact that English has several types of passives such as the *be*-passive, the *get*-passive, and the *have*-

[26] I translated all the quotes from students from Mandarin Chinese.

passive (Celce-Murcia & Larsen-Freeman, 1999), on the cards used in the grammatical difficulty ranking activity,[27] the illustrative sentences (i.e., *The package was delivered to Taiwan*. *Today many products are made of plastic*) represent only the *be-passive* type. In addition, they only represent the passives with simple past and simple present tenses. One participant, Sunny, said that she does not consider the *be*-passive involving simple tense-aspect to be difficult to learn. However, if the passive involves complex tense-aspect, it becomes difficult for her. Similar comments were made by Paula and Joyce. Some students also commented that they perceive learning the use of wh-questions as presented in the illustrative sentences (i.e., "*What is your name?*" and "*Where do you live?*") to be easy. However, they also commented that they were not familiar with other types of wh-questions.

Knowledge of semantics.

Units of information concerning the semantic aspects of a grammar feature were coded as "knowledge of semantics." An example of this is the response, "I do not always understand the cause-effect relationship in a real conditional sentence." Another example is, "I do not fully understand the meanings of the modal auxiliaries."

Features that students perceive as "abstract" were reported to be more difficult to learn. One feature that a couple of students considered abstract is the present perfect. One participant, John, said, "I know that present perfect consists of 'have/has + *p.p.*'" What does *have/has* mean? What does *p.p.* mean? This grammar feature is strange. It is very abstract. I don't get it." Another participant, Fiona, remarked,

> I know that present perfect should be used when we refer to an action that has already taken place. That's my understanding of present perfect. But sometimes in my readings I see that present perfect is used where I don't think it is needed. It confuses me. Present perfect is too abstract and difficult for me to understand.

Features with "multiple" semantic meanings were also reported by some students to be more difficult to learn. For example, a number of students reported that they have difficulty fully understanding the meanings of the modal auxiliaries. They all made comments similar to Kay's remarks:

[27] The illustrative sentences used on the cards are exactly the same as those used in the questionnaire.

> I cannot always distinguish the meanings expressed by different modal verbs. Well, I can tell that *can* and *may* are different in their meanings, but how about *could* and *might*? Do they have the same meaning? I am not sure.

Knowledge of pragmatics.

The responses associated with when and why to use a grammar form were coded as "knowledge of pragmatics." A couple of examples of such responses are like "I don't know when to use the passive,*"* or "I avoid using conditional sentences because they are difficult for me.*"* Many students reported that they are not always clear on when or why to use certain features. Take the passive construction for example. Judy said, "The construction of the passive is simple, but it is not that easy for me to judge when to use an active voice and when to use a passive voice." Likewise, Sandy stated, "The passive is quite easy to learn, but I am not sure when it should be used." A similar comment was also made by Peter, "The passive only has 'auxiliary *be* + P.P.' It should be simple, right? I just don't understand why I sometimes use it incorrectly." Another example is the use of modal auxiliaries. A number of students reported that they are not always clear on when or how to use modal auxiliaries. Beth said, "I am not sure which modal auxiliary I should use. Should I say 'you can come' or 'you may come'? To me, their meanings are quite similar. I am not sure which one I should use." Similarly, Jacky, Tin, Kay, and David also stated that they are not always sure which modal auxiliary best expresses what they mean. Another feature that a number of students reported to be difficult in terms of its use is the article system. Eighteen out of 30 students reported that it is not always clear to them when to add or omit the article "the."

It was also found that learners with inadequate pragmatic knowledge tended to perceive learning of grammar features with "multiple functions" to be easy. For example, one participant, Peter, stated that the articles were not difficult for him to learn. However, he also stated that he knew only two functions of the article system. On the other hand, several students who reported that the articles were not that easy for them to learn seemed to know more about the multiple functions of the article system.

A number of students reported that they tend to avoid using the features that they consider to be difficult. For instance, Jerry said that he avoids using the participial construction in his writing because the feature is too difficult for him to use correctly. David and Nicole also

reported that they avoid using this feature for the same reason. Judy and Stella both stated that they tend to avoid using the present perfect because they do not really understand its use or meaning. Interestingly, however, although many students reported that they perceive clauses to be difficult to learn, they did not report avoiding the use of clauses in their writing. The students reported that they often use clauses in their writing in spite of their difficulty. The primary reason for this is that the students had been told by their teachers that using clauses in their writing makes their writing look "good," thus might benefit them in terms of grades received.

Previous grammar teaching and learning.

Previous grammar teaching and learning refers to the grammar instruction that students received in high school. According to the students, in high school, EFL teachers often taught to tests. That is, the teachers tended to spend more time and effort teaching the features that were most likely tested. In addition, to enhance students' learning of the frequently tested features and to assure that they had learned the features, the teachers often provided the students with a lot of tests or quizzes. The frequently taught and tested features included prepositions, present perfect, clauses, conditional sentences, and participial construction. Several students reported that they consider these "over-tested" features to be more difficult to learn because they sometimes felt overwhelmed and discouraged by the massive number of tests targeting these features. It seemed that the students who said that they did not do well on the tests tend to perceive the tested features to be more difficult to learn, and the students who said they did well on the tests tended to perceive the tested features as less difficult to learn.

Teachers' grammar instruction methods also seemed to influence the students' perceptions of grammatical difficulty. For instance, many students reported that they did not consider real conditionals to be difficult to learn when they were learning only real conditionals. However, after they were taught unreal conditionals, they confused real conditionals with unreal conditionals, and thus, they came to consider real conditionals to be difficult to learn. The reason the students confused real conditionals with unreal conditionals appears to be due to their teachers' instruction of conditional sentences. A number of students said that their teachers took a lot of time and effort to teach the syntactic constituents of unreal conditionals because the teachers thought that these are complex and thus difficult for the students to learn. The students were also provided with a lot of decontextualized practice of the unreal conditionals. However,

the teachers did not provide much instruction on the syntactic constituents of real conditionals because they considered the syntactic constituents of real conditionals to be easy for students. Since more instruction, practice, and testing focused on unreal conditionals rather than on real conditionals, the students were less clear about the syntactic constituents of real conditionals, or they confused real conditionals with unreal conditionals after receiving what they perceived to be unbalanced instruction on the two features. Even though the teachers took much time and effort to teach unreal conditionals, the students still reported that this feature is difficult to learn.

L1 influence.

Some students reported the influence of L1-L2 differences on their learning of certain aspects of grammar features. A few students commented that L1 (Mandarin)-L2 (English) differences adversely influence their understanding the meaning of certain grammar features. For example, Beth commented that Mandarin does not have the present perfect, so this tense is conceptually difficult for her. A few students reported that L1-L2 differences hinder their accurate formulation of certain grammar features. For instance, Abby stated:

> English adjective clauses are used to describe what goes before them. This is different from Mandarin. In Mandarin, we say the adjective clauses first, and then we say the things that the clauses are talking about. In English, it's the opposite. I sometimes forgot that I cannot say adjective clauses first.

Another participant, Fiona, remarked:

> When I want to write a passive sentence, I think about it in Chinese first and then translate it into English. I never write a passive sentence that sounds weird translated from Chinese. But many English passive sentences sound strange to me when I translate them into Chinese. I don't think I write English passive sentences like those strange sentences.

Two points are noteworthy here. First, the excerpts quoted above are mainly from the qualitative data of the ranking results. However, the analysis of the qualitative data from the ranking activity, the stimulated recalls, and the reflective items of the questionnaire generates quite similar themes. In other words, the five themes identified and discussed earlier are applicable across the qualitative datasets obtained in various ways.

Second, although all five factors contribute to learners' perceptions of grammatical difficulty, the "knowledge of syntactic constituents" factor appears most frequently in the learners' remarks. In addition, the data show that learners' perceptions of grammatical difficulty seem to have different levels, and that a learner may perceive the learning of a certain grammar feature to be easy at one level, but not necessarily so at another level. This is illustrated by the following excerpts.

Excerpt 1

Elvira: I think third person *–s* is easy to learn because it has simple rules of thumb. But when I am speaking or writing, I often forgot to add it to the verb when needed.

Excerpt 2

Kay: The passive was very easy for me to learn. You just use "auxiliary *be*" and "p.p." That's it! It's simple. But I am not very clear on when I can use it.

Both Excerpts 1 and 2 indicate that the participants perceive learning of the target features to be easy at the *syntactic* level, but not necessarily so at the *pragmatic* level. Excerpt 3 suggests the opposite.

Excerpt 3

Charlie: Clauses have been difficult for me to learn because there are so many kinds, such as adjective clauses and noun clauses. When I see a clause, I cannot tell which kind of clause it is. However, I use clauses very often in my writing, and I seldom use them incorrectly.

From Excerpts 4 and 5, we see that the participants consider learning of the target features to be difficult at the *semantic* level, but not at the *syntactic* level.

Excerpt 4

Lily: I am not clear about the differences in meaning between *could* and *might* or between *should* and *would*. However, I still consider modal auxiliaries to be easy to learn because they have simple rules. You will get it right if you use the base form of a verb after a modal auxiliary. No exceptions!

Excerpt 5

Thomas: Present perfect is not that easy for me to learn. Present perfect is composed of "have + p.p." The rule is very straightforward. However, it is a bit difficult for me to understand this feature. I don't understand what "have" means.

In sum, learners' perceptions of grammatical difficulty seem to be primarily influenced by the five factors represented by the identified themes. Figure 9 is a schematic display of the five factors influencing learners' perceptions of grammatical difficulty. As Figure 9 indicates, learners' perceptions of grammatical difficulty primarily correlate with their L2 knowledge, L2 grammar learning experience, and L1 knowledge. The dotted lines suggest that learners' perceptions of grammatical difficulty are associated with the five themes indentified.

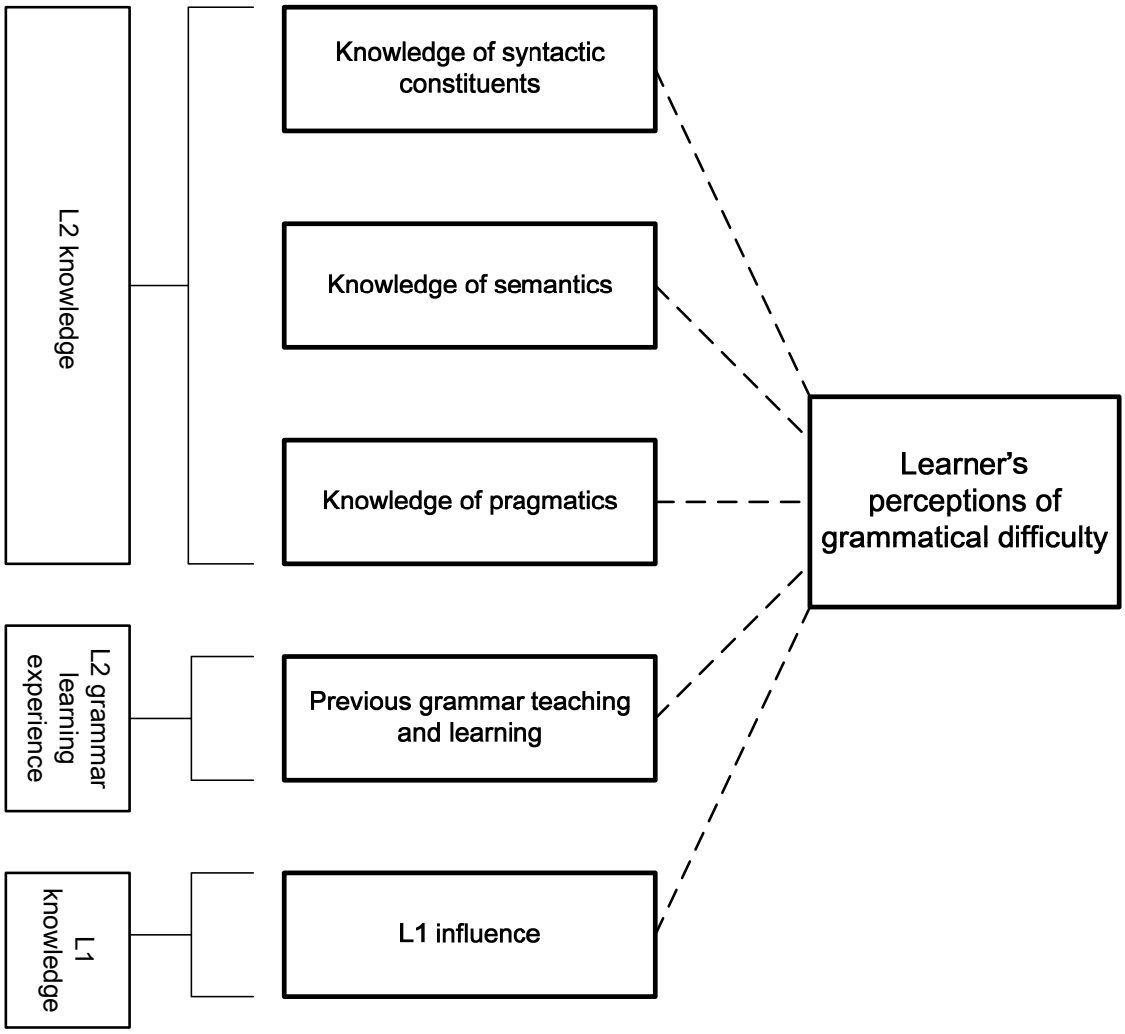


Figure 9. **Factors associated with learners' perceptions of grammatical difficulty.**

Summary

The quantitative data of the questionnaire and the ranking activity were analyzed to address Research Question 1, "Which features of English grammar do university-level Chinese EFL learners perceive as easier, which as more difficult, to learn?" The descriptive statistics of the questionnaire show that overall this group of Chinese EFL learners tend to perceive learning of the 20 target features to be easy. The features were ranked in ascending order based on the value of their mean scores. To validate the questionnaire ranking results, three tests were performed: a Spearman correlation test, a PCA test, and a Pearson product-moment correlation

test. All the test results support the trustworthiness of the grammatical difficulty ranking results from the questionnaire.

The qualitative data gathered from a semi-structured interview, a ranking activity, stimulated recalls, and the reflective items of the questionnaire were analyzed to address Research Question 2, “What influences Chinese EFL learners’ perceptions of grammatical difficulty?” The analysis identified five separate yet intricately interconnected factors associated with Chinese EFL learners’ perceptions of grammatical difficulty. The five factors were “knowledge of syntactic constituents,” “knowledge of semantics,” “knowledge of pragmatics,” “previous grammar teaching and learning,” and “L1 influence.” Although these five factors all contribute to learners’ perceptions of grammatical difficulty, it was observed that the “knowledge of syntactic constituents” factor appears most frequently in learners’ remarks.

Chapter Five
Data Analysis and Results for Research Questions Three and Four

This chapter presents the analyses and results related to Research Questions 3 and 4:

3. Do Chinese EFL learners' perceptions of grammatical difficulty vary according to their general English proficiency level?
4. How do Chinese EFL learners' perceptions of grammatical difficulty relate to their ability to accurately use grammar features on oral production and written metalinguistic tasks?

I begin with a discussion of the validity and reliability testing of the research instruments used to explore the questions. This is followed by a description of the procedures used to analyse the data and the findings.

The Validity and Reliability of the Research Instruments

The research instruments used to explore Research Question 3 include the questionnaire, the ranking activity and the cloze test; those used to explore Research Question 4 include the oral production tasks and metalinguistic task. The validity analyses of the questionnaire and the ranking activity have already been presented in Chapter Four. The cloze test was adopted from Fotos (1991), who demonstrated its validity already. Thus, only the validity of the oral production tasks and the metalinguistic task are discussed here. To establish the validity of the oral production tasks and the metalinguistic task, expert opinion was sought from the members of my thesis committee and my colleagues. Based on their comments and suggestions, the tasks were revised several times before being used in the main study.

The reliability of the cloze test and the other tasks was calculated using Cronbach alpha indexes. The three oral production tasks are the "Package" task, the "Book" task, and the "Travel Agent" task; the first two target the passive construction while the last one targets the real conditional. The metalinguistic task consists of 45 items in total–15 target the passive construction, 15 target the real conditional, and 15 function as distracters. The Cronbach alpha indexes of all the items, the passive items, and the real conditional items were calculated. Table

10 shows the reliability testing results of the cloze test, the three oral production tasks, and the metalinguistic task. As indicated, the reliability levels were high for the cloze test, the "Book" task, the "Travel Agent" task, and the items targeting the real conditional: α = .84, .89, .87, and .86, respectively. However, they were a bit low for the "Package" task and the items targeting the passive construction, α = .63 and .61, respectively. The reliability of the "Package" task is lower than expected; it might be because only 5 items are subject to the analysis in this task. The low reliability for the items targeting the passive construction may be related to students' good performance and low variation in scores. Although the reliabilities of these two tasks are low, they still fall within an acceptable range. Thus, a decision was made to include the data from these two tasks for subsequent analysis.

Table 10

Reliability of the Cloze Test, the Oral Production Tasks, and the Metalinguistic Task

	Test/Task	Cronbach alpha Index
Cloze Test		.84
Oral production tasks	The "Package" task	.63
	The "Book" task	.89
	The "Travel Agent" task	.87
Metalinguistic task	All 45 items	.78
	Items targeting the passive construction	.61
	Items targeting the real conditional	.86

Data Analysis for Research Question 3

Research Question 3 explores whether Chinese EFL learners' perceptions of grammatical difficulty vary according to their English proficiency level. Learners' perceptions of grammatical difficulty were explored by means of a questionnaire; their general English proficiency was assessed through the administration of a cloze test. For analysis of the questionnaire, see Chapter Four. The cloze test was completed by 277 students. Unlike Fotos (1991), who used an "exact-word" scoring method, I used an "appropriate-word" scoring method. The exact-word scoring method gives points only for answers that are exactly the same as the prepared answer keys, whereas the appropriate-word method also gives points for acceptable alternatives to the prepared answers. For example, in the sentences *We could not keep three fast-growing lions forever. Sadly, we decided that the two big ones had to go*, the underlined words (*not*, *sadly*, and *ones*) are on the answer key. For the response *We could not keep three fast-growing lions forever. Therefore, we decided that the two big lions had to go*, the exact-word scoring method generates 1 point because only one response matches the prepared answer key, whereas the use of the appropriate-word scoring method generates 3 points because the answers *therefore* and *lions* are acceptable alternatives. The appropriate-word scoring method was employed for two reasons. First, scores obtained by this method more accurately indicate learners' overall L2 proficiency (Heilenman, 1983; Laesch & van Kleeck, 1987). Second, appropriate-word scoring is less likely to generate a narrow range of scores (Fotos, 1991); and a narrow range of scores might render the test less reliable (Bachman, 2004). The maximum score for the test is 50, one point for each correct, i.e., acceptable, answer.

Descriptive statistics of the cloze test were computed in order to get an overview of the data. Students' responses to each of the 20 closed-ended items in the questionnaire were analyzed against the proficiency test scores using a Spearman correlation test. The Spearman correlation test was employed in light of the absence of a normal distribution in the learners' responses to the questionnaire. Using the Bonferroni approach to control for Type I errors across the 20 correlations, a p value of less than .0025 (.05/20 = .0025) was required for significance.

To further explore whether learners' perceptions of grammatical difficulty differ according to their general proficiency level, the qualitative data of 10 participants gathered during the ranking activity were examined. The 10 participants are a sub-sample of the 30

students who completed Phase 2 of the research.[28] The 10 participants were selected based on their scores on the proficiency test; among the 30 participants, 5 participants with the lowest scores and 5 participants with the highest scores were chosen for this analysis. These two groups were created in order to have students representing two distinctively different proficiency levels. It will be recalled that the qualitative data generated five factors that are associated with learners' perceptions of grammatical difficulty: *knowledge of syntactic constituents*, *knowledge of semantics*, *knowledge of pragmatics*, *previous grammar teaching and learning,* and *L1 influence* (see Chapter Four for discussion of these five factors). The data from the 10 participants were re-examined to search for instances of these five factors. The data were coded by a fellow doctoral candidate and me, with a reliability coefficient $k = .89$. Disagreements between the coders were resolved through discussion. For each group, the frequency of the instances of each of the five factors was calculated. The summed frequency counts from the two groups were then compared. The content of the instances was also examined and compared qualitatively.

Results for Research Question 3

The descriptive statistics of the cloze test show a wide, normal distribution of the scores, with a minimum of 2 and a maximum of 36 (M= 16.82, SD = 6.97), suggesting that the participants' proficiency levels span a wide range. Figure 10 shows the distribution of the test scores of the 277 participants.

[28] Phase 2 of the research involved a semi-structured interview, a ranking activity, two cloze activities, and two stimulated recalls. However, in this analysis, only the data gathered during the ranking activity were used.

***Figure 10*. Distribution of the scores of the cloze test completed by 277 participants.**

Table 11 presents the results of the Spearman correlation test that examined the relationship between the students' perceptions of difficulty for 20 grammatical features as measured by the questionnaire and their proficiency scores as measured by the cloze test. In Table 11, the 20 features are presented in ascending order on the basis of their mean scores.

Table 11

Spearman Correlation Coefficients between Students' Perceptions of the Difficulty of the 20 Features and their Cloze Test Scores (N = 277)

Ranking order	Items	Cloze test
1	Negation	-.16
2	Third person *–s*	-.11
3	Present progressive	-.09
4	Simple past *–ed*	.10

Ranking order	Items	Cloze test
5	Wh-questions	.10
6	Modal auxiliaries	.15
7	Adjective comparatives	-.04
8	Articles	.10
9	Passives	-.13
10	Past progressive	-.21*
11	(Un)countable noun	.03
12	Present perfect	-.04
13	Question tags	-.26*
14	Infinitives	-.22*
15	Clauses	-.11
16	Embedded questions	-.18*
17	Prepositions	-.09
18	Real conditionals	-.22*
19	Participial construction	-.23*
20	Unreal conditionals	-.18

Note. * $p < .0025$ (2-tailed).

As Table 11 shows, all the correlation coefficients are low, and only six of them are statistically significant, suggesting weak correlations between ranking of learners' perceptions of the difficulty of individual grammar features and their general English proficiency level.

Table 12 presents the frequencies of the five factors by the two groups differing in proficiency (i.e., high proficiency vs. low proficiency). These five factors were identified from the qualitative data collected in Phase 2 of the research, which explored what might influence learners' perceptions of grammatical difficulty: *knowledge of syntactic constituents*, *knowledge of semantics*, *knowledge of pragmatics*, *L1 influence*, and *previous grammar teaching and*

learning. As Table 12 indicates, the two proficiency groups are quite similar in terms of the frequency of the five factors.

Table 12

Frequency Counts of the Five Themes from the Two Proficiency Groups (n=10)

	Knowledge of syntactic constituents	Knowledge of semantics	Knowledge of pragmatics	Previous grammar teaching and learning	L1 influence	Total
High proficiency	30 (38%)	6 (8%)	24 (31%)	14 (18%)	4 (5%)	78
Low proficiency	30 (39%)	10 (13%)	21 (27%)	12 (16%)	4 (5%)	77
Total	60 (39%)	16 (10%)	45 (29%)	26 (17%)	8 (5%)	155

With regard to the content analysis, the qualitative analysis showed that the high proficiency group and the low proficiency group were quite similar in terms of the content for all the factors, except for "knowledge of pragmatics." The references to knowledge of pragmatics made by the low proficiency learners primarily concern their difficulty with or uncertainty about certain features. Comments such as "I often use this feature inaccurately" or "I don't know when or how to use the feature" predominate. The high proficiency group generated more comments related to the use of the features and these were varied; some were about the ease with which they use certain features, some were about the difficulty of using certain features, and some were about how they used certain features.

Data Analysis for Research Question 4

Research Question 4 explores how Chinese EFL learners' perceptions of grammatical difficulty relate to the accuracy of their use of grammar features on different tasks. Learners' use of the target features was elicited through the administration of three oral production tasks and the metalinguistic task. As previously noted, the target features for the oral production and metalinguistic tasks are passive construction and real conditionals. A Wilcoxon test to examine the questionnaire responses of the 26 participants who participated in the oral production and metalinguistic tasks indicated that their responses to the items representing these two features significantly differed with respect to their level of difficulty, $z = -3.99$, $p < .01$, $r = -.78$. The results suggest that overall the students perceived the passive construction to be easier to learn than the real conditional.

What follows is a discussion of the analysis of the learners' performance on the oral production and metalinguistic tasks.

Oral production tasks.

The students' performance on the three tasks was recorded and then transcribed following the conventions listed in Appendix M. The transcribed data were then scored for accuracy in the manner illustrated in Table 13. The examples are taken from the database for this study.

Table 13

Scoring Schemes for the Oral Production Tasks and Illustrative Examples

Tasks	Scoring rules	Scores	Examples
The OPTs targeting the passive construction (i.e., the "Package" task and the "Book" task)	Correct construction of passive	2	• *The package was sent to Toronto USA..* • *The pyramids were built in 2700 BC.*
	Malformed passive • Past participle was malformed	1	• *The package was deliver to Toronto USA..* • *The pyramids were builded in 2700 BC.*

Tasks	Scoring rules	Scores	Examples
	• Subject-verb agreement was violated		• *The pyramids was built in 2700 BC.*
	Incorrect construction of passive	0	• *The package deliver to Toronto USA..* • *The pyramids built in 2700 BC.*
	No attempt to use the passive	0	• *Someone sent the package.* • *People built the pyramids.*
	Items left unanswered	0	
OPT targeting the real conditional (i.e., the "Travel Agent" task)	Correct construction of real conditionals	2	• *If you stay at the River Hotel, you will pay only eighty dollars per night.*
	Correct use of the verb in either the if-clause or the main clause	1	• *If you will stay at the River Hotel, you pay only eighty dollars per night.* • *If you stay at the River Hotel, you had been saved only eighty dollars per night.*
	Incorrect use of the verbs in both the if-clause and the main clause	0	• *If you stays at the River Hotel, you pays only eighty dollars per night.*
	No attempt to use the real conditional	0	• *If you were a member, you would have got a fifteen percent discount on flights.*
	Items left unanswered	0	

As indicated in Table 13, students were awarded one point for their "partially" successful attempts to form a passive construction or the verb of a real conditional sentence. This is based on the assumption that students' knowledge of the target features is not all-or-nothing, and thus their partially successful attempt should be acknowledged. Using the scoring schemes described, the maximum total scores for the "Package" task, the "Book" task, and the "Travel Agent" task are 10, 10, and 18, respectively. Given the differences in the maximum scores for each measure, percentages were used in statistical analyses. For the testing of the inter-rater reliability, the three oral production tasks were scored by another doctoral candidate and me. The inter-rater reliability of the three oral production tasks is high, r = .99, .99, and .93, $p < .05$, for the "Package" task, the "Book" task, and the "Travel Agent" task, respectively.

Next, learners' perceptions of grammatical difficulty were explored in relation to their ability to accurately use the target features that they identified as more or less difficult to learn. The performance of 26 students on the oral production tasks targeting the passive construction (i.e., the "Package" task and the "Book" task) was compared with their performance on the oral task targeting the real conditional (i.e., "Travel Agent" task) using Wilcoxon tests. A non-parametric analysis was used because of the absence of a normal distribution of the learners' responses. Using the Bonferroni approach to control for Type I errors across the 2 correlations, a p value of less than .025 (.05/2 = .025) was required for significance.

Written metalinguistic task.

As previously noted, in total there are 45 items in the metalinguistic task, which consists of three parts: error identification, correction, and explanation. The scoring schemes for these three parts are described and illustrated in Table 14. The examples are taken from the database of this study.

Table 14

Scoring Schemes for the Metalinguistic Task and Illustrative Examples

Part	Scoring principles	Score	Examples
The passive construction			
Error identification	Correct identification	1	*Laws are make by the government.* ➢ The ungrammatical part is make.
	Incorrect identification	0	*Laws are make by the government.* ➢ The ungrammatical part is are.
	Item left unanswered	0	
Correction	Correct form is provided	2	*Laws are make by the government.* ➢ The correct form is made.
	Partially successful attempt to give the correct form • Malformed passive is provided • Subject verb agreement is not followed	1	*Laws are make by the government.* ➢ The correct form is maked. The houses on this street built last year. ➢ The correct form is was built.
	Incorrect form is provided	0	*Laws are make by the government.* ➢ The correct form is were.

Part	Scoring principles	Score	Examples
	Item left unanswered	0	
Explanation	Provided an accurate explanation of the rule in question	2	*Laws are make by the government.* ➢ It is ungrammatical because "Passive construction = aux be + p.p. [past participle]." *Many messages were receiving yesterday.* ➢ It is ungrammatical because "The subject is the receiver of the action denoted by the verb. Therefore, it is the passive voice rather than the active voice that should be used here".[29]
	Provided a partially correct or less than satisfactory explanation of the rule in question	1	*Laws are make by the government.* ➢ It is ungrammatical because "the past participle of the verb *make* is *maked*." *Vitamins were discovering in the early 1900's.* ➢ It is ungrammatical because "*V-ing* cannot be used in a passive sentence."
	Provided an incorrect, irrelevant, or non explanation of the rule in question	0	*Many passages were receiving yesterday.* ➢ It is ungrammatical because "past tense should be used here." *His heart has been breaking three times this year.*

[29] All students' explanations were mostly written in Mandarin Chinese. They are translated by the researcher.

Part	Scoring principles	Score	Examples
			➢ It is ungrammatical because "*broken* is used in present perfect." ➢ It is ungrammatical because "it sounds strange."
	Item left unanswered	0	
The real conditional			
Error identification	Correct identification	1	*If he taking the train, he will get to work by 8:00.* ➢ The ungrammatical part is taking.
	Incorrect identification	0	*If he taking the train, he will get to work by 8:00.* ➢ The ungrammatical part is will.
	Item left unanswered	0	
Correction	Correct form is provided	2	*If he taking the train, he will get to work by 8:00.* ➢ The correct form is takes.
	Partially successful attempt to give the correct form • Subject verb agreement is not	1	*If he taking the train, he will get to work by 8:00.* ➢ The correct form is take. *If my parents will come to visit me, I will take*

Part	Scoring principles	Score	Examples
	followed		*them to Taipei.* ➢ The correct form is comes.
	Incorrect form is provided	0	*If he taking the train, he will get to work by 8:00.* ➢ The correct form is got.
	Item left unanswered	0	
Explanation	Provided an accurate explanation of the rule in question	2	*If my parents will come to visit me, I will take them to Taipei.* ➢ It is ungrammatical because "simple present tense should be used in the *if* clause of a future conditional sentence." *If children eat breakfast, they having more energy.* ➢ It is ungrammatical because "present tense (or future tense) should be used in the main clause of a present (or future) real conditional sentence."
	Provided a partially correct or less than satisfactory explanation of the rule in question	1	*If I using the computer too much, my eyes will hurt.* ➢ It is ungrammatical because "*V-ing* cannot be used in the verb part of a conditional sentence." ➢ It is ungrammatical because it is a

Part	Scoring principles	Score	Examples
			conditional sentence. *If children eat breakfast, they having more energy.* ➢ It is ungrammatical because "it should use future tense."
	Provided an incorrect, irrelevant, or non explanation of the rule in question	0	*If he taking the train, he will get to work by 8:00.* ➢ It is ungrammatical because "*be* verb is needed here." *If children eat breakfast, they having more energy.* ➢ It is ungrammatical because "it sounds strange."
	Item left unanswered	0	

Using the scoring schemes described in Table 14, the maximum scores for the identification, correction, and explanation are 1, 2, and 2, respectively, and the maximum total score for each item is 5. The maximum scores for the items targeting the passive construction and those targeting the real conditional are 75 (5 points x 15 items) each. Percentages are used in the statistical analyses. As noted in Table 14, incorrect answers and items left unanswered receive 0 points. As both items receive 0, I computed and reported the frequency of items left unanswered in the three sections in order to know exactly how many were incorrect answers versus items left unanswered. To examine inter-rater reliability, the metalinguistic task was scored by an experienced EFL teacher and me. The inter-rater reliability is high, r = .98, .99,

and .98, $p < .05$, respectively for (1) all the items, (2) the items targeting the passive construction, and (3) the items targeting the real conditional.

To explore learners' perceptions of grammatical difficulty in relation to their ability to accurately use the target features that they identified as more or less difficult to learn in the completion of the metalinguistic task, a paired *t* test was performed to compare the learners'[30] overall performance on the items targeting passives and real conditionals. A series of Wilcoxon tests were conducted to compare learners' performance on each of the three parts (i.e., error identification, correction, and explanation) of the tasks. The choice of using a paired *t* test or a Wilcoxon test depends on the presence or absence of normal distribution in the relevant data, results of which are presented in the Results section. To control for Type I errors across comparisons of the three parts, the Bonferroni approach was used, indicating a *p* value of less than .02 (.05/3 = .02) required for significance.

Results for Research Question 4

Table 15 presents descriptive statistics for the three oral production tasks and the metalinguistic task. As indicated in the table, the mean percentages for the three oral production tasks are 60.38, 75, and 86.32 respectively, while the mean percentages for the metalinguistic task are 89.38 and 80.51, respectively for the items targeting the passive construction and those targeting the real conditional. With regard to the standard deviation of the oral production tasks, the "Book" task has the largest standard deviation, while the "Travel Agent" task has the smallest. The standard deviation for the items targeting the passive construction ($SD = 9.82$) is smaller than that for the items targeting the real conditional ($SD = 16.35$). The information regarding the skewness and kurtosis indicates that the "Book" task and the "Travel Agent" task do not have normal distribution, while the others are all normally distributed.

[30] In total, 185 participants completed the metalinguistic task. However, to address Research Question 4, only the data of the 26 participants who completed both the oral production tasks and the metalinguistic task were analyzed.

Table 15

Descriptive Statistics of the Oral Production Tasks and the Metalinguistic Task (n = 26)

	Minimum (%)	Maximum (%)	*M* (%)	*SD*	Skewness		Kurtosis	
					Statistic	Std. error	Statistic	Std. error
OPT (the "Package" task)	.00	100.00	60.38	27.35	-.50	.46	-.36	.89
OPT (the "Book" task)	.00	100.00	75.00	30.76	-1.55	.46	1.52	.89
OPT (the "Travel Agent" task)	22.22	100.00	86.32	20.51	-2.25	.46	5.02	.89
Metalinguistic task (items targeting the passive construction)	68.00	100.00	89.38	9.82	-.88	.46	-.04	.89
Metalinguistic task (items targeting the real conditional)	36.00	100.00	80.51	16.35	-.88	.46	.51	.89

Table 16 provides the descriptive statistics for the three parts (error identification, correction, and explanation) of the metalinguistic task. As indicated in the table, for both the items targeting the passive construction and the items targeting the real conditional, "explanation" has the smallest mean percentage while "error identification" has the largest mean percentage. Likewise, the explanation has the largest standard deviation, while the error correction has the smallest standard deviation. The information regarding skewness and kurtosis indicates absence of normal distribution on all of the parts except the explanation of the real conditional.

Table 16

Descriptive Statistics of the Three Parts of the Metalinguistic Task

	Minimum (%)	Maximum (%)	*M* (%)	*SD*	Skewness		Kurtosis	
					Statistic	Std. error	Statistic	Std. error
Items targeting the passive construction								
Error identification	73.33	100.00	94.62	7.31	-1.4	.46	1.57	.89
Correction	63.33	100.00	88.21	10.92	-.98	.46	.13	.89
Explanation	66.67	100.00	87.95	11.93	-.53	.46	-1.24	.89
Items targeting the real conditional								
Error identification	73.33	100.00	94.10	8.07	-1.56	.46	1.80	.89
Correction	30.00	100.00	85.64	16.70	-1.74	.46	3.74	.89
Explanation	20.00	100.00	68.59	23.84	-.39	.46	-.78	.89

Overall, there were 390 responses to each of the three parts of the metalinguistic task (15 items x 26 participants). Table 17 shows the sum of the frequency of the items left unanswered in each part. The percentage provided is obtained by a calculation of the number of the items left unanswered divided by 390. As can be seen in Table 17, very few items were left unanswered.

Table 17

Frequency of the Items Left Unanswered in the Metalinguistic Task

Passive Construction			Real Conditionals		
Identification	Correction	Explanation	Identification	Correction	Explanation
4	5	7	2	1	5
1%	1%	2%	0.5%	0.2%	1.2%

Wilcoxon tests were conducted to compare learners' performance on the two oral production tasks targeting the passive construction and the one targeting the real conditional. As both the "Package" task and the "Book" task target the passive construction and the "Travel Agent" task targets the real conditional, the students' performance on the two tasks was compared with that on the "Travel Agent" task. A comparison of the results indicates no significant difference between the students' performance on the "Book" task and the "Travel Agent" task (z = -1.32, p = .19). However, there is a significant difference between their performance on the "Package" task and the "Travel Agent" task, z = -3.06, p = .002, r = -0.60; learners' performance on the "Travel Agent" task targeting the real conditional was better than their performance on the "Package" task targeting the passive construction.

Analyses of the metalinguistic task included a paired t test to explore learners' overall performance on the items targeting the passive construction and those targeting the real conditional. A t test was used because, as indicated in Table 15, the learners' overall performance on the items targeting the passive construction and their performance on the items targeting the real conditional were both normally distributed. The results showed that the mean

for the "passive" items (M = 89.38, SD = 9.82) is significantly greater than the mean for the "real conditional" items (M = 80.51, SD = 16.35), $t(25) = 2.46$, $p = .02$, with a medium effect size, $d = .48$. Learners' performance on the three individual parts of the passive items was compared with those on the real conditional items, using Wilcoxon tests. The results indicate that, on average, there is no significant difference between their performance on the two error identification parts (z= -.06, $p = .95$) or the two error correction parts (z = -.37, $p = .71$). However, their performance on the explanation part of the passive items was significantly better than on the real conditional items, z =–3.04, $p = .002$, r =-.60.

Summary

To address Research Question 3, which explores whether Chinese EFL learners' perceptions of grammatical difficulty vary according to their general English proficiency level, the students' responses to each of the 20 items in the questionnaire were analyzed against the proficiency test scores using a Spearman correlation test. Results indicate weak and for the most part insignificant correlations. The analyses of the responses of the high-proficiency and the low-proficiency learners to further explore whether proficiency level is related to learners' perceptions of grammatical difficulty indicate that these two groups are quite similar in terms of the frequency of the five factors associated with their perceptions of grammatical difficulty. However, the results of the content analysis indicate that these two groups differ somewhat at the pragmatic level in terms of their content.

To address Research Question 4, which investigates how learners' perceptions of grammatical difficulty relate to their L2 task-based performance, learners' performance on the three oral production tasks were compared using Wilcoxon tests, and their performance on the items targeting the passive construction in the metalinguistic task was compared with their performance on the items targeting the real conditional, using a paired t test and Wilcoxon tests. Results show that the learners performed significantly better on the "Travel Agent" task (i.e., the task targeting the real conditional) than on the "Package" task (one task targeting the passive construction). However, there is no significant difference between their performance on the "Travel Agent" task and on the "Book" task (the other task targeting the passive construction). A comparison of their performance on the metalinguistic task shows that learners performed better on the items targeting the passive construction than on the items targeting the real conditional.

With regard to their performance on the three individual parts of the metalinguistic task, results show that there is no significant difference between the error identification and error correction parts. However, learners performed significantly better on the explanation part for the passive items than for the real conditional items.

The results related to the four Research Questions investigated in this study reported in Chapters 4 and 5 will now be discussed in Chapter 6.

Chapter Six
Discussion and Conclusions

This chapter discusses the results relative to the four Research Questions, identifies limitations of the study, considers the implications of the findings, and provides suggestions for future research.

Discussion

Research Question 1.

The first Research Question asks which features of English grammar university-level Chinese EFL learners perceive as easier, and which as more difficult, to learn. The question was explored by the administration of a questionnaire, which was designed to explore learners' perceptions of grammatical difficulty of 20 target features. The mean scores of the 20 items on the questionnaire cluster toward the lower end of the scale; the lowest mean score is 1.17 and the highest is 3.32. This positively skewed response distribution suggests that, on the whole, this sample of Chinese EFL learners did not perceive any of the 20 target features to be very difficult, although some of the features were perceived to be more difficult than the others. This finding might be, in part, due to the fact that the questionnaire explores difficulty primarily at the syntactic level, and English learning at this level tended to be perceived as relatively easy by this group of learners because of their years of de-contextualized, form-oriented L2 learning experience. In the questionnaire, the target features are presented via the use of parts of speech and illustrative sentences in which they are highlighted (e.g., I *have finished the job*.) This format emphasizes the syntactic aspect of the target features, which might have predisposed the learners to reflect upon their prior learning of that aspect of the target features. In this research context, English language teaching in high school often focuses on the development of grammatical knowledge about English. This is supported by the students' reports in the questionnaire, which showed that grammar is frequently taught in EFL instruction in senior and junior high school.[31] The qualitative data of the present study also confirm this observation; many students

[31] In the questionnaire, the students were asked how often they were taught grammar in junior and senior high school. Results show that 31.9 % of the students indicated "very often", 46.5% indicated "often" in junior high and 50.5 % indicated "very often" and 38.8 % indicated "often" in senior high.

participating in this phase of research reported that a considerable portion of the English instruction they received in high school was allotted to the teaching of English grammar. The qualitative data further suggest that although the functions of grammar features are taught in these learning contexts, especially in senior high school, much of the focus is on the syntactic constituents of the grammar features. The syntactic constituents are often presented in overly simplified "formulas," for example, "*be* + V-ing" (symbolizing the syntactic constitution of the present progressive) or "*be* + P.P. (past participle)" (symbolizing the passive construction). Basic rules of thumb are frequently used to help students learn the formation of the syntactic constituents. Here "rule of thumb" refers to overly simplified grammar rules used to help learners to learn the formation of grammar structures. To illustrate, a rule of thumb used to teach the syntactic constituents of English simple past might be something like, "to form a simple past tense, add *–ed* to the base form of the verb." Given their many years of learning syntactic constituents and their associated rules of thumb, the learners considered them to be relatively easy. This speculation is supported by the results of the qualitative data: many students participating in this phase of research reported that, overall, they perceived learning of the "grammar rules" of the target features to be easy. Since the questionnaire explored difficulty at the syntactic level and the learners' prior L2 learning tends to focus on this aspect, this might explain why this group of Chinese EFL learners tended to perceive learning of the target features to be easy.

Although the learners tended to perceive the learning of the 20 target features to be easy, some of the features were perceived to be easier than others. The six features ranked as "least difficult" in the questionnaire are *negation*, *third person –s*, *present progressive*, *simple past –ed*, *wh-question*s, and *modal auxiliaries*, while the six features ranked as "most difficult" are *clauses*, *embedded questions*, *prepositions*, *real conditionals*, *participial construction*, and *unreal conditionals*.[32] An examination of the features in these two groups suggests that they can be distinguished from each other by the extent of the metalanguage needed to formulate a basic rule of thumb. As noted earlier, for this group of students, a basic rule of thumb is frequently used to help them learn the formulation of the target features. Consequently, students might base their judgments of grammatical difficulty on the complexity of the rule they have been taught.

[32] As previously noted, because the descriptive statistics of the questionnaire show that the mean score differences for some features are small, only the six features with the lowest mean scores and the six features with the highest mean scores are discussed here.

Pedagogical rules for the formation of a target structure can be more or less complex depending on the degree of elaboration with which the target structure is formulated (Housen, et al., 2005). A basic rule of thumb limits the elaboration to a minimum, yet still provides sufficient information to explain the formulation of the target feature.

In the "least difficult" group, all of the features, except for wh-questions, can be formulated comparatively simply. To illustrate, the basic rule of thumb for the formulation of the present progressive might be something like, "to form a present progressive verb, use *be* plus *V-ing*." However, the formulation of wh-questions needs more extensive use of metalanguage than that needed for the formulation of the other five features in the "least difficult" group. One reason why these learners perceived wh-questions to be easy to learn might be that their perceptions were biased by the sample sentences used for this feature. As revealed in the qualitative data gathered during the ranking activity, a couple of students reported that they considered the two sample sentences— "*What is your name?*" and "*Where do you live?*"—to be easy to learn because these two sentences are not only taught early in EFL instruction but are also frequently taught as "chunks."

If we examine those features in the "most difficult" group, all of the features, except prepositions, require considerably more use of metalanguage in the rules of thumb that explain their formulation. For example, the basic rule of thumb for *real conditionals* is something like, "to form real conditionals, write an *if*-clause and a result clause. Use present tense in the if-clause, and use present/future tense or modal auxiliaries plus base form of the verb in the result clause." Even more metalanguage is required to explain the formulation of "tenses" and "clauses." Prepositions, on the other hand, do not require as much use of metalanguage because in terms of their syntactic constituents (e.g., preposition + noun phrase) they are comparatively simple. However, as revealed in the qualitative results, the students tended to consider prepositions to be difficult to learn because, in their opinion, the learning of this feature depends largely on memorization as there are few useful basic rules of thumb to explain their use. In short, the questionnaire ranking results suggest that, generally speaking, learners' perceptions of grammatical difficulty are associated with the amount of metalanguage needed to formulate a basic rule of thumb for the features in question; the less metalanguage needed for their formulation, the less difficult they are to learn as perceived by this group of learners, and vice versa.

However, the questionnaire ranking findings of the present study are not shared by Scheffler's (2009) study, which, to my knowledge, is the only study that has investigated learners' perceptions of grammatical difficulty using a similar approach. Scheffler investigated the effectiveness of L2 instruction in relation to grammatical difficulty from the perspective of Polish EFL learners. A 5-point Likert scale questionnaire was administered to 50 advanced learners to explore their perceptions of the learning difficulty of 11 target features. Based on their mean scores, the 11 features are ranked below from least to most difficult.

1. adjectives and adverbs
2. pronouns
3. nouns
4. articles
5. passive voice
6. reported speech
7. conditional sentences
8. modal verbs
9. *–ing* forms and infinitives
10. prepositions
11. tenses

Comparing Scheffler's results with those of the present study, both groups of EFL learners tended to perceive prepositions to be more difficult to learn, and articles and passive voice to be less difficult to learn. However, whereas the Polish EFL learners perceived tenses to be the most difficult to learn, the Chinese EFL learners considered conditional sentences to be the most difficult to learn. The discrepancy might be in part due to the difference in the learners' L1 (Mandarin Chinese vs. Polish). The influence of L1 on L2 learning has been extensively documented in the literature (e.g., Lightbown & Spada, 2000; Odlin, 1989, 2003; Spada & Lightbown, 1999; L. White, 1991). Thus, it seems reasonable to suppose that the difference in the learners' L1 might impact their perceptions of grammatical difficulty. However, the fact that there are no other studies of this kind with which to compare the results of these two, it remains unclear whether the ranking differences are due to L1 influence (or other factors). More empirical investigation into learners' perceptions of grammatical difficulty is warranted.

Research Question 2.

The second Research Question asks: What influences Chinese EFL learners' perceptions of grammatical difficulty? Analysis of the qualitative data indicates that Chinese EFL learners' perceptions of grammatical difficulty are primarily influenced by five distinguishable, yet intricately interconnected factors: *knowledge of syntactic constituents*, *knowledge of semantics*, *knowledge of pragmatics*, *previous grammar teaching and learning*, and *L1 influence*. More detailed findings pertaining to the five factors are summarized below.

- Learners' perceptions of grammatical difficulty can be examined with reference to the syntactic, semantic, and/or pragmatic levels. It is not unusual for a learner to consider learning of a certain grammar feature to be easy at one level, but not at another.
- At the syntactic level, learners tend to perceive grammar features that can be formulated based on simple grammatical rules of thumb easier to learn than those that need complex grammatical rules of thumb. However, although some of the target features were perceived to be more difficult to learn than others in terms of grammar rules, many of the students participating in this phase of research also commented that overall they perceived learning of the 10 target features selected for the ranking activity to be relatively easy.
- At the semantic level, learners tend to perceive grammar features with more abstract meaning (e.g., *present perfect*) to be more difficult to learn than those with less abstract meaning.
- At the pragmatic level, many students who had little difficulty verbalizing syntactic rules reported that they have difficulty using the features correctly. In addition, many students reported that they tend to avoid using grammar features that they consider difficult to learn.
- Overall, among the five factors, "knowledge of syntactic constituents" appears most frequently in the students' discussions of grammatical difficulty.

- It was observed that students' test performance on the grammar features and their impression of their teachers' grammar instruction exert some influence on their perceptions of grammatical difficulty.
- Some students are aware of the influence of L1-L2 differences on their learning of different aspects of grammar features; for example, their understanding of the meaning of certain L2 grammar features (e.g., present perfect), their formulation of certain L2 grammar features (e.g., adjective clauses), or their use of certain L2 grammar features (e.g., passive).

The finding that learners' perceptions of grammatical difficulty can be examined with reference to the syntactic, semantic, and pragmatic levels supports Larsen-Freeman's (2001 and elsewhere) argument that grammatical difficulty should be examined in relation to *form*, *meaning*, and *use*. As noted earlier, according to Larsen-Freeman, "form" concerns the accurate use of the (morpho)syntactic aspect of a grammar feature. "Meaning" is primarily concerned with understanding the inherent or literal message encoded by a lexical item or a lexico-grammatical feature. "Use" concerns using a lexico-grammatical feature appropriately within a context. Larsen-Freeman argues that a grammar feature can be easy with respect to one aspect, but difficult with respect to another. The finding supports her argument. The finding also suggests that the learners in the current study were conscious of the different levels of learning difficulty when dealing with different aspects of L2 grammar.

It was observed that at the syntactic level, the learners tend to perceive grammar features that need simple grammatical rules of thumb for their formulation to be easier to learn than those that need complex grammatical rules of thumb. This qualitative finding is in line with, and thus, supports, the previously discussed interpretation of the ranking results of the questionnaire. Some researchers (e.g., Housen et al., 2005; Hulstijn, 1995; Hulstijn & de Graaff, 1994) claim that there is a positive correlation between grammatical difficulty and the inherent complexity of rules. The finding of the current study suggests that from the perspective of L2 learners, the positive correlation between the two holds. However, the findings also indicate that many students perceived the rules governing most of the target features to be easy. This resonates with the findings obtained from the questionnaire, supporting the assumption that the questionnaire findings are primarily concerned with the syntax of the target features.

The reason that *knowledge of syntactic constituents* appears most frequently in the students' discussions of grammatical difficulty might be because their previous form-oriented grammar learning experiences made them more likely to draw on this domain of knowledge than on semantic or pragmatic knowledge in their discussion of grammatical difficulty. In addition, the fact that the research instruments used emphasized the syntactic aspect of the target features made it less likely that students would consider other domains of language use or knowledge in their discussion.

At the semantic level, the finding that learners tended to perceive grammar features with "abstract" meaning to be more difficult to learn corresponds to DeKeyser's (2005) claim that grammatical difficulty can sometimes be partially attributed to problems of *meaning*. At the pragmatic level, many learners reported that they tend to avoid using grammar features that they consider difficult to learn. This finding resonates with the observation made by several SLA researchers (Hinkel, 2004; Hulstijn & Marchena, 1989; Kleinmann, 1978; Schachter, 1974) that L2 learners do indeed avoid using grammar features that are difficult for them. The results also indicate that many students who had little difficulty articulating syntactic rules are aware that they have difficulty using them correctly. This finding resonates with the claim that some features are easy to describe, but difficult to produce (DeKeyser & Sokalski, 1996).

It is interesting that learners' test performance may also exert some influence on their perceptions of grammatical difficulty. Situated in a learning context where test performance is highly valued, it is not surprising that the learners judged ease or difficulty of learning target features by their test performance. Many students reported that in high school EFL teachers often taught to tests. This suggests the existence of "washback"[33] in the EFL learning context in which the study was conducted. In addition, the finding that learners' impression of their teachers' instructional methods contributes to learners' perceptions of grammatical difficulty suggests that L2 teachers play an important role in determining whether learners perceive the learning of grammar to be easy or not. Loewen et al. (2009) report an association between L2 learners' previous grammar learning experience and their attitudes toward grammar learning and error correction. The present study further demonstrates that L2 learners' previous grammar learning experience affects their perceptions of grammatical difficulty. That is, students who considered

[33] "Washback" refers to, "the extent to which the instruction and use of a test influences language teachers and learners to do things they would not otherwise do that promote or inhibit language learning" (Messick, 1996, p. 241).

their teachers' instructional methods not to be particularly helpful for learning certain grammar features tended to perceive those features as difficult, and vice versa.

The influence of L1 (Chinese) on L2 (English) learning has been extensively documented in the literature (e.g., Chou, 2000; Hinkel, 2002; McEnery & Xiao, 2005; Yip, 1995). The current study found that some learners are aware that the L1-L2 differences may somewhat adversely influence their learning of *form*, *meaning* or *use* of certain L2 grammar features. For example, a couple of learners reported that the L1-L2 differences negatively affect their learning of formulation of adjective clauses. Some learners stated that the differences hinder their understanding of the meaning of the present perfect, and some commented that the differences affect their use of the passives. Given the comments, it seems reasonable to suppose that grammatical awareness associated with L1 influence may somewhat impact learners' perceptions of grammatical difficulty of certain grammar features. However, to what extent and in what way it may influence learners' perceptions of grammatical difficulty merit further empirical investigation.

Research Question 3.

The third Research Question asks whether Chinese EFL learners' perceptions of grammatical difficulty vary according to their general English proficiency level. The Spearman correlation test results reveal low and mostly insignificant correlations between learners' perceptions of the learning difficulty of individual grammar features (explored by the questionnaire) and their proficiency levels (operationalized as their cloze test scores). In the current study, learners' overall English proficiency was assessed by means of a cloze test, which arguably measures learners' integrative ability to use the target language (Fotos, 1991; Hanaia & Shikhani, 1986). Thus, the low correlations suggest that learners' judgments of learning difficulty at the syntactic level are not related to their integrative ability to use the target language. Given that the participants perceived learning difficulty primarily in terms of syntactic constituents and rules of thumb, it is not surprising that they perceived a feature that is easy to describe (e.g., third person *–s*) to be easy to learn, and a feature that is difficult to describe (e.g., participial construction) to be difficult to learn regardless of their proficiency level.

To further explore whether learners' perceptions of grammatical difficulty vary according to their proficiency, the frequency with which the five factors of grammatical difficulty emerged from the learners in the high and low proficiency groups were analyzed and compared. The five factors are *knowledge of syntactic constituents*, *knowledge of semantics*, *knowledge of pragmatics*, *previous grammar teaching and learning*, and *L1 influence*. Results show that the two proficiency groups were quite similar in terms of the frequency of the five factors. For both groups, on average, the factor that appeared the most frequently was *knowledge of syntactic constituents*. This finding gives some statistical support for the claim made on the basis of the qualitative analysis of the data that knowledge of syntactic constituents predominates in the learners' discussion of grammatical difficulty. As discussed previously, in light of their prior grammar-oriented English learning experience, it is no surprise that the learners, regardless of their proficiency level, frequently referred to the syntactic constituents of the target features in their discussion of grammatical difficulty. The similarity in learners' language learning experiences may also explain why the two groups were similar with respect to the frequency of the factor *previous grammar teaching and learning*.

However, the finding that the groups were similar in terms of how frequently *knowledge of pragmatics* was given as an explanation for difficulty is somewhat unanticipated. I had assumed that the high proficiency group would refer to knowledge of pragmatics more often than the low proficiency group because the former presumably had more experience using the features than the latter. However, the results of the content analysis of the five factors help to explain this unexpected finding. The results show that the two groups are similar in terms of content for all the factors, except for "knowledge of pragmatics." The references to knowledge of pragmatics made by the low proficiency learners primarily concerned their difficulty with or uncertainty about certain features. Comments such as "I often use this feature inaccurately" or "I don't know when or how to use the feature" predominate. The high proficiency group produced more mixed comments; some were about the ease with which they use certain features, some were about the difficulty of using certain features, and some were about how they used certain features. Overall, the qualitative findings suggest that at the pragmatic level, learners' perceptions of grammatical difficulty may vary according to proficiency levels. However, this suggestion is at most tentative and is subject to the caveat that the sample size that generated this data was small (n = 10).

Research Question 4.

The fourth Research Question asks how Chinese EFL learners' perceptions of grammatical difficulty relate to their ability to accurately use grammar features on oral production and written metalinguistic tasks. The findings related to this Research Question are summarized below.

- Two features—*passive construction* and *real conditionals*—were selected as the target features for the oral production and the metalinguistic tasks. The questionnaire results indicate that the 26 learners who participated in this phase of the research perceived the passive construction to be easier to learn than the real conditional.
- The learners performed significantly better on the "Travel Agent" task, which targets the real conditional, than on the "Package" task, which targets the passive construction. However, there was no significant difference between their performance on the "Travel Agent" task and on the "Book" task (the other task targeting the passive construction).
- With regard to the learners' performance on the metalinguistic task, results indicate that in general they performed significantly better on the items targeting the passive construction than on the items targeting the real conditional. In terms of their performance on the three individual parts of the metalinguistic task (error identification, correction, and explanation), the learners performed best on the error identification part, less well on the correction part, and least well on the explanation part. In addition, there was no significant difference between their performance on the "error identification" or "correction" parts of the tasks targeting the passive construction and the real conditional. However, the learners performed significantly better on the explanation part of the passive construction task than on the real conditional task.
- In terms of the relationship between learners' perceptions of grammatical difficulty and their actual use of the two target features on the oral production tasks, results show that the feature perceived to be less difficult to learn (i.e., the passive construction) was not used more accurately by this sample group of learners than the feature perceived to be more difficult to learn (i.e., the real conditional), whereas on

the metalinguistic task, the feature perceived to be less difficult to learn was used more accurately.

What follows first considers the learners' metalinguistic task performance in relation to their perceptions of grammatical difficulty of the target features. This is followed by a discussion of the findings regarding the relationship between their oral production task performance and their perceptions of grammatical difficulty.

With regard to learners' metalinguistic task performance, they performed significantly better on the items targeting the passive construction than on the items targeting the real conditional. However, the findings regarding their performance on the three individual parts of the task further show that the difference is only in the "explanation" part. This may be in part due to the fact that it is easier to render the rule of thumb governing the passive construction than those governing the real conditional because describing the formulation of a passive sentence requires less extensive use of metalanguage than describing the formulation of a real conditional sentence. To illustrate, the rule of thumb governing the passive construction might be given as, "To form a passive, move the object of a sentence to the subject position, change the verb to aux be + past participle, and then put the word "by" in front of the object." On the other hand, the rule of thumb governing the real conditional requires more extensive use of metalanguage as it entails not only a description of how to form if-clauses and result clauses, but also the use of verb tenses in both clauses. Consequently, the findings obtained from the metalinguistic task are congruent with the findings from the questionnaire in that they both can be explained by differences in the amount of metalanguage required to formulate the rule for the target features. Similar to the questionnaire results indicating that grammar features governed by rules that can be described using less metalanguage are considered to be less difficult to learn, on the metalinguistic task, the feature that could be described more simply was described better. These findings taken together suggest that a feature with simpler rules is perceived as easier to learn and is used more accurately on the task that requires learners to elaborate upon their knowledge of the features in question.

In terms of type of L2 knowledge, arguably, the findings from the comparison of the learners' perceptions of grammatical difficulty and their metalinguistic task performance seem mainly applicable to explicit knowledge, specifically explicit knowledge of the syntax of the

target features. Given their previous grammar-oriented learning experiences, the learners who participated in this phase of the research were assumed to possess at least some explicit knowledge of the target grammar features. This assumption is supported by qualitative data showing that these learners were able to articulate partial rules of thumb for many of the grammar features in question.[34] Other research (e.g., Elder et al., 1999; Roehr, 2007; White & Ranta, 2002) also indicates that form-focused instruction can increase a learner's explicit knowledge and promote its use. Accordingly, it seems reasonable to suppose that the learners participating in the questionnaire survey had some explicit knowledge of the target features at their disposal. Researchers (Alderson, Clapham, & Steel, 1997; Roehr, 2008; Sorace, 1985) postulate that tasks that focus on the language form itself are conducive to the retrieval of *explicit knowledge* of the features in question. In light of this postulation, it is assumed that the use of metalinguistic terminology in the questionnaire may have primed the learners to rely extensively on their explicit knowledge of the target features to judge grammatical difficulty. In the current study, "explicit knowledge" refers to conscious, declarative knowledge that is acquired either through a process of conscious reflection on associated implicit knowledge or through the formal study of grammar in the L2 classroom (Roehr, 2008).[35] Likewise, learners' metalinguistic task performance can be seen as indicative of the use of their explicit knowledge of target features. Therefore, the claim that a feature that requires less extensive use of metalanguage to describe its formulation would be perceived as easier to learn and used more accurately on tasks that focus on language form seems mainly applicable at the level of explicit knowledge and specifically to the syntactic aspect of target features. R. Ellis (2006b) argues that at the explicit level, grammar rules that can be formulated with less extensive use of metalanguage are easier to learn than those that need more extensive use of metalanguage. The findings with regard to the metalinguistic task performance support his argument.

The findings, however, do not support the claim that learners would perceive a grammar feature as less difficult to learn if they had better explicit knowledge of it. First, the findings obtained from the metalinguistic task do not provide a clear answer to the question of whether the learners had better explicit knowledge of the passive construction than of the real conditional.

---

[34] Nineteen out of 26 students who participated in this part of the study also participated in phase 2 of the research in which the qualitative data were collected.

[35] Roehr (2008) and others (e.g., Alderson et al. 1997; Elder & Manwaring, 2004) refer to explicit knowledge, as defined here, as "metalinguistic knowledge."

Learners performed similarly on the "error identification" and "error correction" sections for both features and the only significant difference was in the "explanation" section, which arguably may be mainly due to the nature of the structures investigated. As well, several SLA researchers have cautioned that verbalization of rules is not a precise measure of explicit knowledge and that learners vary in their ability to verbalize rules (Bialystok & Ryan, 1985; Ellis, 1991, 2005; Han & Ellis, 1998; Hu, 2002; Macaro & Masterman, 2006). Accordingly, we cannot be certain whether the learners had different levels of explicit knowledge of the two features.

With regard to the learners' oral production task performance, the passive construction was not used more accurately than the real conditional. However, these results should be taken with caution since the three oral production tasks differ in terms of their task structure, which may have affected the learners' performance. In terms of task structure, among the three tasks, the "Travel Agent" task (the task targeting the real conditional) has the most constraints; it specifically requires the learners to start their sentences with "If you.., you…." Such a requirement is likely to orient learners' attention to language form rather than meaning. In contrast, the "Package" task (one of the tasks targeting the passive construction) is the least controlled; it has relatively few constraints, and thus provides a context that is conducive to the spontaneous use of the target feature. Compared with the "Travel Agent" task, the "Package" task is more likely to orient learners' attention to meaning rather than form. The "Book" task (the other task targeting the passive construction) has fewer constraints than the "Travel Agent" task, but more than the "Package" task. The finding that the learners performed better on the "Book" task (mean accuracy percentage = 75.00) than on the "Package" task (mean accuracy percentage = 60.38) suggests that the two tasks are not parallel in terms of their requirements, which in part may be due to the differences in their task structure. Furthermore, although it would have been better to have had two oral production tasks targeting the real conditional, this was not possible due to the difficulty of creating oral production tasks to elicit this feature and none appropriate for this study could be found in the existing literature. Thus, because differences in task structures may have confounded learners' performance (e.g., Doughty, 2003; Yuan & R. Ellis, 2003) and there was an unequal number of oral production tasks targeting the two target features, the finding that the learners did not use the passive construction more accurately than the real conditional needs to be taken with caution. That is, the finding that the feature perceived to be less difficult to learn was not used more accurately remains tentative.

In terms of type of L2 knowledge, it is believed that, in general, oral tasks in which learners are asked to spontaneously produce language within time constraints encourage greater use of implicit knowledge (R. Ellis, 2005). Although the three oral production tasks used in this study are not strictly timed tasks, because of the nature of one-on-one interaction, learners were under pressure to produce language. Thus, the findings regarding the oral production task performance might be viewed as indicative of a greater use of the learners' implicit knowledge of target features. Implicit knowledge here refers to unconscious linguistic knowledge that cannot be verbalized on demand (Hulstijn, 2005; R. Ellis, 2005). The finding that the feature perceived to be less difficult to learn was not used more accurately on the oral production tasks might be because learners' perceptions of grammatical difficulty explored by the questionnaire are more closely related to learners' explicit knowledge, whereas the oral task performance is more closely related to learners' implicit knowledge. Learners' implicit and explicit knowledge of the grammar features do not necessarily match (Krashen, 1982; R. Ellis, 2006b, 2008). Moreover, R. Ellis (2006b) argues that a feature can be easy to learn as explicit knowledge, but not as implicit, and vice versa. In accordance with R. Ellis' argument, another speculation is that the passive construction may be less difficult to learn explicitly, but more difficult to learn implicitly, while the opposite might be true for the real conditional. This speculation is supported by the findings of the current study, which show that the learners used the passive construction more accurately on the metalinguistic task than on the oral production tasks, while they used the real conditional more accurately on the oral production task than on the metalinguistic task (see Table 15 in Chapter Five). However, this speculation warrants further empirical verification.

Conclusions and Implications

This exploratory study set out to investigate the issue of grammatical difficulty from L2 learners' perspectives in relation to their overall L2 proficiency and L2 performance and knowledge. The findings of the current study indicate that Chinese EFL learners' perceptions of grammatical difficulty are influenced by several factors related to their L2 knowledge, L2 grammar learning experience, and L1 knowledge. Furthermore, while learners' perceptions of grammatical difficulty included reference to syntactic, semantic, and/or pragmatic levels, overall, this group of Chinese EFL learners perceived learning of the target features at the syntactic level to be relatively easy. Notwithstanding, learners perceived syntactic features that require more

extensive use of metalanguage to describe their formulation as more difficult to learn than those that can be described more simply.

In terms of the relationship between learners' perceptions of grammatical difficulty and their overall L2 proficiency, the current study found that learners' perceptions of grammatical difficulty do not vary according to their overall L2 proficiency at the syntactic level, but they may vary at the pragmatic level. With regard to the relationship between learners' perceptions of grammatical difficulty and their metalinguistic task performance, results suggest that the feature perceived to be less difficult to learn is used more accurately on the task that focuses on the language form itself. In terms of type of L2 knowledge, the findings arguably are mainly applicable to explicit knowledge, specifically the syntactic aspect of the target features. With regard to the relationship between learners' perceptions of grammatical difficulty and their oral production task performance, results show that the feature perceived to be less difficult to learn is not necessarily used more accurately. However, due to the possibility that the learners' performance may have been affected by the type of oral production tasks used, the findings should be taken with caution. Notwithstanding, in terms of L2 knowledge, the findings suggest that learners' perceptions of grammatical difficulty based on explicit knowledge of the target features may not match task performance that relies on greater use of the associated implicit knowledge.

A number of implications can be drawn from the findings of the current study. First, the findings raise questions about the role of learners' prior L2 learning experience in determining their perceptions of grammatical difficulty. The present study found learners' grammar-oriented learning experience to have a great influence on their perceptions of grammatical difficulty. It would be interesting to explore how learners with different learning experience (for example, "meaning-based") would perceive grammatical difficulty and how their perceptions of grammatical difficulty would relate to their overall L2 proficiency and L2 performance and knowledge.

Second, the findings with regard to learners' task performance have implications for the measurement of L2 knowledge. It was found that the learners performed significantly better on the "Book" task than on the "Package" task, $z = -2.42$, $p = .015$, $r = -.47$, which both targeted the passive construction. Although both tasks ask learners to respond to a picture strip, whereas the

"Package" task asks the learners to narrate a story, using the target verbs provided in the pictures, the "Book" task asks the learners to use the provided target verbs in conjunction with the provided temporal information to describe some non-interrelated historical facts. As the "Book" task has more constraints than the "Package" task, the "Book" task may be more likely to direct learners' attention to language form than the "Package" task, which consequently may involve greater use of explicit knowledge. Thus, the finding supports researchers' (e.g., Bialystok & Ryan, 1985) argument that oral production tasks may involve the use of explicit knowledge to varying degrees, depending on the task-related factors. The finding also suggests that "task structure" plays an essential role in determining the extent of implicit/explicit knowledge used in completing the tasks. In light of the finding, task structure should be taken into consideration when evaluating the results of tasks used to measure learners' implicit knowledge of grammar features. Furthermore, in terms of learners' metalinguistic task performance, it was found that the learners performed similarly on the "error identification" parts and the "error correction" parts, but they performed significantly differently on the "explanation" parts. The finding confirms the findings of previous research (Alderson, et al., 1997; Elder & Manwaring, 2004; Green & Hecht, 1992; Sorace, 1985) that learners can correct errors instantiating the rule in question even though they are unable to state the rule. However, the differential performance on the explanation part might be in part due to the structural nature of the target features—that is, the formulation of one feature is easier to describe than the other. As such, the finding raises questions about the role of the structural nature of a grammar feature in mediating learners' expression of their explicit knowledge of the feature. This further raises the question of the extent to which one can infer learners' explicit knowledge of a target feature by their ability to correct errors and to explain their corrections. In short, the findings regarding learners' task performance confirm that measuring learners' L2 knowledge is methodologically challenging, and highlight the need for using multiple tests to get a more complete picture of L2 learners' linguistic knowledge (Chaudron, 2003; DeKeyser, 2003; Doughty, 2003; Norris & Ortega, 2003).

Third, methodologically, the present study demonstrates the advantages of using mixed methods allowing for multi-level analysis of complex issues and increasing the validity of research outcomes (Dornyei, 2007; Greene, Kreider, & Mayer, 2005). The qualitative data collected in the present study contribute to an in-depth understanding of the research problem investigated. They also help to interpret statistical results of the study. In addition, that the

quantitative findings based on the questionnaire and the qualitative components of the research have corresponding findings not only enhances our confidence in the validity of the questionnaire findings, but also increases the generalizability of the qualitative findings.

Fourth, from a pedagogical perspective, the findings of the study provide several implications that are useful for L2 teachers in general and EFL instructors teaching in this research context (i.e., in Taiwan) in particular. The finding that learners perceive grammatical difficulty on different levels (syntactic, semantic, and pragmatic) suggests that when they describe features as easy or difficult to learn, it is essential to be aware of the level of learning difficulty that they are referring to. Also as the findings from this study suggest, learners who receive instruction that is primarily focused on grammar are less likely to develop a sufficient awareness and understanding of other levels of difficulty (e.g., semantic, pragmatic, etc.). Therefore, it is important for teachers to draw learners' attention to all aspects of the challenges of learning language forms (e.g., syntactic, semantic, and pragmatic aspects). This may be particularly important in instructional approaches that focus exclusively on one aspect of language or another (e.g., form and meaning). In addition, as revealed in the qualitative data, some students reported that they tend to avoid using the features they perceive to be difficult to learn. Thus, knowing which features the students tend to perceive to be difficult to learn would help the teachers to be aware which features the students might be avoiding in their language use. Furthermore, as the findings of the current study suggest, for the students whose prior language experience is primarily form-oriented, neither their ability to verbalize the grammar rule nor their perceptions of grammatical difficulty is a reliable predictor of their ability to use grammar features in communicative tasks. The ease/difficulty of describing the formulation of a feature is not a reliable predictor, either. Therefore, L2 instruction should include opportunities for the communicative use of grammar features that learners have no difficulty verbalizing their rules, and that they perceive as easy to learn or to describe.

Limitations of the Study

The findings of this study must be interpreted in light of several limitations with regard to the research design and the research instruments used. First, the current study was conducted in an EFL context where the learners' L1 is primarily Mandarin and their L2 learning is primarily grammar-oriented. The participants were all first-year university level EFL students. The

findings of the study, therefore, may not be applicable to other L2 learner populations. In addition, the findings cannot be generalized to grammar features other than those investigated in the study. Second, the number (n = 26) of participants for the oral production tasks and the number (n = 10) of the participants included for the frequency analysis of the five themes are small. Therefore, the relevant findings are subject to this caveat. An increase in the number of participants would strengthen the robustness and reliability of the findings.

Third, the questionnaire used in the current study has two disadvantages. The format (i.e., the use of parts of speech and illustrative sentences) of the questionnaire items emphasizes the syntactic aspect of the target features. Thus, the questionnaire findings are more informative about learners' perceptions of grammatical difficulty at the syntactic level, but less informative about the pragmatic and semantic levels. Another disadvantage is related to the use of two short, simple sentences to exemplify the use of the target features. The qualitative findings indicated that learners perceived the use of the target features illustrated by the sample sentences in the card ranking activity to be easy to learn. Since the parts of speech and the illustrative sentences used in the cards are exactly the same as those used in the questionnaire, and there is a high correlation (r = .86) between the two, it seems reasonable to suppose that the questionnaire findings are restricted by the illustrative sentences used in the questionnaire, too.

Fourth, in the current study, the choice of the cloze test as an L2 proficiency test was made for pragmatic reasons, i.e., for its brevity of administration. Although there is sufficient evidence to support its validity in predicting learners' general proficiency (Fotos, 1991), the cloze test used is by no means an ideal measure of L2 proficiency (Brown, 1983, 1989; Hanania & Shikhani, 1986). The fact that listening and speaking components are not included in this test limits its ability to produce a full assessment of learners' ability with, and knowledge of the target language.

Two additional limitations relate to the oral production tasks used. As noted earlier, it would have been preferable to have administered two oral production tasks targeting the real conditional. Furthermore, there was a lack of equivalency across the three oral production tasks; they were not parallel in terms of their structure. This makes comparisons of the learners' performance across the tasks problematic. To circumvent the task equivalency problem, perhaps

a better choice of test would have been an oral "elicited imitation test" (Erlam, 2006) as tests of this kind are more comparable in terms of their overall structure and requirements.

Directions for Future Research

To my knowledge, the present study is the first empirical attempt to explore L2 learners' perceptions of grammatical difficulty in relation to their L2 proficiency, L2 performance and L2 knowledge. As noted earlier, the findings of the current study may not be applicable to other L2 learner populations, and the findings cannot be generalized to grammar features other than those investigated. Accordingly, replication studies in a variety of learning contexts (for example, EFL, ESL, immersion programs), with learners of different age groups and L1s, and with different target features are warranted. Replication studies could also respond to some of the limitations of the current study with regard to the research instruments. For example, in light of the disadvantages of the questionnaire described, the validity of the questionnaire as a tool for exploring learners' perceptions of grammatical difficulty would be improved by (1) discarding the use of metalinguistic terminology, (2) including as many types of the target features as possible in the illustrative sentences, and (3) in the questionnaire instructions, instead of asking learners whether they perceive the target features to be easy to "learn," ask them whether they perceive the target features to be easy to "use." The last revision aims to more specifically direct learners to reflect upon their experience of using the grammar features rather than their experience of learning their syntactic constituents and rules of formulation. As well, in light of the limitations noted of using oral picture-cued production tasks, perhaps a better choice is the use of an oral elicited imitation test to elicit the spontaneous use of the grammar features in question. This kind of test would enable us to compare learners' spontaneous use of the target grammar features across the tasks.

In addition to replication studies, it would be of interest to explore learners' perceptions of grammatical difficulty in relation to their language learning aptitude, specifically their language analytic ability. Previous research has found positive correlations between learners' language analytic ability and their L2 learning (Harley & Hart, 1997a; Ranta, 2002; Sheen, 2007a). Following this, it would be of interest to see whether learners' perceptions of grammatical difficulty vary according to individual differences in their language analytic ability.

Another fruitful investigation would be a comparison between learners' and teachers' perceptions of grammatical difficulty. Like research on learners' perceptions of grammatical difficulty, research on teachers' perceptions of grammatical difficulty is also thin on the ground. Research on L2 teacher education indicates that most L2 teachers consider explicit knowledge of grammar to be desirable and beneficial for their learners (e.g., Schulz, 1996, 2001). Furthermore, given that teachers' classroom practices are influenced by their beliefs, based on prior language learning experience and professional training (Borg, 2003), it seems reasonable that their perceptions of grammatical difficulty might inform their instructional decision-making. Nevertheless, to date, there is a paucity of research exploring the issue of grammatical difficulty from the perspective of L2 teachers. There is also little research comparing learners' and teachers' perceptions of grammatical difficulty. These under-researched areas warrant further investigation. The need for a comparison of teachers' and students' perceptions of grammatical difficulty is further supported by the observation that the teachers with whom I shared the findings of the questionnaire were all surprised that the students, overall, tended to perceive learning of the target features to be easy. The teachers' reaction suggests that they perceive grammatical difficulty differently from the students. This speculation, however, needs empirical verification.

In conclusion, this exploratory study of learners' perceptions of grammatical difficulty in relation to their overall L2 proficiency and L2 performance and knowledge opens up a potentially productive area of research. I hope that the pursuit of this line of inquiry will provide useful information for L2 researchers, L2 teachers and L2 learners.

References

Alderson, J., Clapham, C., & Steel, D. (1997). Metalinguistic knowledge, language aptitude and language proficiency. *Language Teaching Research, 1*(2), 93-121.

Ammar, A. (2008). Prompts and recasts: Differential effects on second language morphosyntax *Language Teaching Research, 12*(2), 183-210.

Ammar, A., & Lightbown, P. M. (2005). Teaching marked linguistic structure-more about the acquisition of relative clauses by Arab learners of English. In A. Housen & M. Pierrard (Eds.), *Investigation in instructed second language acquisition* (pp. 167-198). New York, NY: Mouton de Gruyter.

Ammar, A., & Spada, N. (2006). One size fits all?: Recasts, prompts, and L2 learning. *Studies in Second Language Acquisition, 28*(4), 543-574.

Anderson, J. (1976). *Language, memory and thought*. Mahwah, NJ: Erlbaum.

Anderson, J. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. (1990). *Cognitive psychology and its implications*. New York, NY: Freeman.

Anderson, J. (1993). *Rules of the mind*. Hillsdale, NJ: Lawrence Erlbaum.

Anderson, J. (2000). *Learning and memory: An integrated approach* (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Anderson, J. (2005). *Cognitive psychology and its implications* (6th ed.). New York, NY: Worth.

Anderson, J., Bothell, D., Douglass, B., Lebiere, C., & Qin, Y. (2004). An integrated theory of mind. *Psychological Review, 111*(4), 1036-1060.

Anderson, J., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum.

Anderson, R., & Shirai, Y. (1996). The primacy of aspect in first and second language acquisition: The pidgin/creole connection. In W. Ritchie & T. Bhatia (Eds.), *Handbook on language acquisition* (pp. 527-571). New York, NY: Academic Press.

Bachman, L. (1996). *Fundamental considerations in language testing*. Oxford, England: Oxford University Press.

Bachman, L. (2004). *Statistical analyses for language assessment*. Cambridge, England: Cambridge University Press.

Bard, E. G., Robertson, D., & Sorace, A. (1996). Magnitude estimation of linguistic acceptability. *Language, 72*, 32-68.

Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning and use*. Malden, MA: Blackwell.

Berent, G. (1985). Markedness considerations in the acquisition of conditional sentences. *Language Learning, 35*(3), 337-372.

Berry, R. (2005). Making the most of metalanguage. *Language Awareness, 14*, 3-20.

Berry, R. (2009). EFL major's knowledge of metalinguistic terminology: A comparative study. *Language Awareness, 18*(2), 113-128.

Bialystok, E. (1978). A theoretical model of second language learning *Language Learning, 28*, 69-83.

Bialystok, E. (1979). Explicit and implicit judgements of L2 grammaticality *Language Learning, 29*, 81-103.

Bialystok, E. (1981). The role of linguistic knowledge in second language use. *Studies in Second Language Acquisition, 4*(1), 31-45.

Bialystok, E. (1991). Metalinguistic dimensions of bilingual language proficiency. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 113-140). New York, NY: Cambridge University Press.

Bialystok, E. (1994). Analysis and control in the development of second language proficiency. *Studies in Second Language Acquisition, 16*, 157-168.

Bialystok, E., & Ryan, E. B. (1985). A metacognitive framework for the development of first and second language skills. In D. Forrest-Pressley, C. MacKinnon & T. Waller (Eds.), *Metacognition, cognition, and human performance* (Vol. 1) (pp. 207-252). Orlando, FL: Academic Press.

Birdsong, D. (1989). *Metalinguistic performance and interlinguistic competence*. Berlin, Germany: Springer-Verlag.

Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teacher think, know, believe, and do. *Language Teaching, 36*, 81-109.

Brown, J. (1983). A closer look at cloze: Validity and reliability. In J. Oller (Ed.), *Issues in language testing research* (pp. 237-250). Rowley, MA: Newbury House.

Brown, J. (1989). Cloze item difficulty. *JALT Journal, 11*, 46-67.

Brown, J., & Rodgers, T. S. (2002). *Doing second language research*. Oxford, England: Oxford University Press.

Brown, J. D. (2001). *Using surveys in language programs*. Cambridge, MA: Cambridge University Press.

Carroll, J. B., & Sapon, S. (1959). *The modern language aptitude test*. San Antonio, TX: Psychological Measurement.

Celce-Murcia, M., & Larsen-Freeman, D. (1999). *The grammar book: An ESL/EFL teacher's course*. Boston, MA: Heinle & Heinle.

Chaudron, C. (1983). Research in metalinguistic judgement: A review of theory, methods and results. *Language Learning, 33*, 343-377.

Chaudron, C. (2003). Data collection in SLA research. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 762-823). Oxford, England: Blackwell.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chou, C. (2000). Chinese speakers' acquisition of English conditionals: Acquisition order and L1 transfer effects. *Second Language Studies, 19*(1), 57-98.

Collins, L., Trofimovich, P., White, J., Cardoso, W., & Horst, M. (2009). Some input on the easy/difficult grammar question: An empirical study. *Modern Language Journal, 93*, 336-353.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd. ed.). Hillsdale, NJ: Lawrence Erlbaum.

Davies, W., & Kaplan, T. (1998). Native speaker vs. L2 learner grammaticality judgments. *Applied Linguistics, 19*, 183-203.

de Graaff, R. (1997). The eXperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition, 19*, 249-276.

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 42-63). New York, NY: Cambridge University Press.

DeKeyser, R. (2001). Automaticity and automatization. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 125-151). Cambridge: Cambridge University Press.

DeKeyser, R. (2003). Implicit and explicit learning In C. J. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 313-348). Malden, MA: Blackwell.

DeKeyser, R. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning, 55*(S1), 1-25.

DeKeyser, R. (2007). Introduction: Situating the concept of practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 1-18). New York, NY: Cambridge University Press.

DeKeyser, R., Salaberry, R., Robinson, P., & Harrington, M. (2002). What gets processed in processing instruction? A commentary on Bill VanPatten's "Processing Instruction: An update". *Language Learning, 52*(4), 805-823.

DeKeyser, R., & Sokalski, K. (1996). The differential role of comprehension and production practice. *Language Learning, 46*, 613-642.

Devitt, S. (2000). Book Review: Language Processing and Second Language Development: Processability Theory. *System, 28*, 455-460.

Dienes, Z., & Perner, J. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences, 22*, 735-808.

Dornyei, Z. (2003). *Questionnaire in second language research: Construction, administration, and processing*. New York, NY: Lawrence Erlbaum.

Dornyei, Z. (2007). *Research methods in applied linguistics*. Oxford, UK: Oxford University Press.

Dornyei, Z. (2009). *The psychology of second language acquisition*. Oxford, England: Oxford University Press.

Doughty, C. (1991). Second language instruction does make a difference: Evidence from an empirical study of SL relativization. *Studies in Second Language Acquisition, 13*, 431-469.

Doughty, C. (2003). Instructed SLA: Constraints, compensation, and enhancement. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 256-310). New York, NY: Blackwell.

Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 114-138). Cambridge, England: Cambridge University Press.

Doughty, C., & Williams, J. (1998a). Pedagogical choices in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197-262). Cambridge, England: Cambridge University Press.

Doughty, C., & Williams, J. (Eds.). (1998b). *Focus on form in classroom second language acquisition*. Cambridge, England: Cambridge University Press.

Dulay, H., & Burt, M. (1974). Natural sequences in child second language acquisition. *Language Learning, 24*, 37-53.

Dulay, H., Burt, M., & Krashen, S. (1982). *Language two*. Oxford, England: Oxford University Press.

Ehrman, M. E., & Oxford, R. (1995). Cognition plus: Correlates of language learning success. *The Modern Language Journal, 79*, 67-89.

Elder, C., & Manwaring, D. (2004). The relationship between metalinguistic knowledge and learning outcomes among undergraduate students of Chinese. *Language Awareness, 13*(3), 145-162.

Ellis, N. (1994). Introduction: Implicit and explicit language learning--An overview. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 1-31). London, England: Academic Press.

Ellis, N. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143-188.

Ellis, N. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition, 27*, 305-352.

Ellis, R. (1990). *Instructed second language acquisition*. Oxford, England: Blackwell.

Ellis, R. (1991). Grammaticality judgements and second language acquisition. *Studies in Second Language Acquisition, 13,* 161-186.

Ellis, R. (1994a). *The study of second language acquisition*. Oxford, England: Oxford University Press.

Ellis, R. (1994b). A theory of instructed second language acquisition *Implicit and explicit learning of languages* (pp. 79-114). London, England: Academic Press.

Ellis, R. (2001). Introduction: Investigating form-focused instruction. *Language Learning, 51*(S1), 1-46.

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: A psychometric study. *Studies in Second Language Acquisition, 27*, 141-172.

Ellis, R. (2006a). Current issues in the teaching of grammar: An SLA perspective. *TESOL Quarterly, 40*(1), 83-108.

Ellis, R. (2006b). Modelling learning difficulty and second language proficiency: The differential contributions of implicit and explicit knowledge. *Applied Linguistics, 27*(3), 431-463.

Ellis, R. (2007). The differential effects of corrective feedback on two grammatical structures. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 339-360). Oxford, England: Oxford University Press.

Ellis, R. (2008). Investigating grammatical difficulty in second language learning: Implications for second language acquisition research and language testing. *International Journal of Applied Linguistics, 18*(1), 4-22.

Ellis, R. (2009a). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 3-25). Tonawanda, NY: Multilingual Matters.

Ellis, R. (2009b). *The study of second language acquisition* (2nd ed.). Oxford, England: Oxford University Press.

Ellis, R., Loewen, S., & Basturkmen, H. (2006). Disentangling focus on form. A response to Sheen and O'Neill (2005). *Applied Linguistics, 27*(1), 135-141.

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research, 9*(2), 147-171.

Farley, A. P. (2004). *Structured input: Grammar instruction for the acquisition-oriented classroom*. New York, NY: McGraw-Hill.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). Thousand Oaks, CA: Sage.

Fotos, S. (1991). The cloze test as an integrative measure of EFL proficiency: A substitute for essays on college entrance examinations? *Language Learning, 41*(3), 313-336.

Fotos, S. (2005). Traditional and grammar translation methods for second language teaching. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 653-670). Mahwah, NJ: Lawrence Erlbaum.

Gass, S. (1997). *Input, interaction and the second language learner.* Mahwah, NJ: Lawrence Erlbaum

Gass, S., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum.

Gass, S., & Mackey, A. (2002). Frequency effects and second language acquisition: A complex picture? *Studies in Second Language Acquisition, 24*, 249-260.

Gass, S., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd. ed.). New York, NY: Routledge.

Gass, S., & Selinker, L. (Eds.). (1992). *Language transfer in language learning*. Amsterdam: John Benjamins.

Gass, S., Svetics, I., & Lemelin, S. (2003). Differential effects of attention. *Language Learning, 53*(3), 497-545.

Gillham, B. (2007). *Developing a questionnaire* (2nd ed.). New York, NY: Continuum.

Givon, T. (1991). Markedness in grammar: Distributional, communicative and cognitive correlates of syntactic structure. *Studies in Language 15*(2), 335-370.

Givon, T. (1995). *Functionalism and grammar.* Amsterdam: John Benjamins.

Goldschneider, J. M., & DeKeyser, R. M. (2005). Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning, 55*(S1), 27-77.

Goss, N., Zhang, Y., & Lantolf, J. (1994). Two heads may be better than one: Mental activity in second-language grammaticality judgments. In E. Tarone, S. Gass & A. Cohen (Eds.), *Research methodology in second-language acquisition* (pp. 263-286). Hillsdale, NJ: Lawrence Erlbaum.

Green, P., & Hecht, K. (1992). Implicit and explicit grammar: An empirical study. *Applied Linguistics, 13*(2), 168-184.

Greene, J. C., Kreider, H., & Mayer, E. (2005). Combining qualitative and quantitative methods in social inquiry. In B. Somekh & C. Lewin (Eds.), *Research methods in the social sciences* (pp. 274-282). London, UK: Sage.

Han, Y., & Ellis, R. (1998). Implicit knowledge, explicit knowledge and general language proficiency. *Language Teaching Research, 2*, 1-23.

Han, Z. (2000). Persistence of the implicit influences on NL: The case of the pseudo-passive. *Applied Linguistics, 21*(1), 55-82.

Hanania, E., & Shikhani, M. (1986). Interrelationships among three tests of language proficiency: Standardized ESL, cloze and writing. *TESOL Quarterly, 20*, 97-109.

Harley, B., & Hart, D. (1997a). Language aptitude and second-language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition, 19*, 379-400.

Harley, B., & Hart, D. (1997b). Language aptitude and second language proficiency in classroom learners of different starting ages. *Studies in Second Language Acquisition, 19*(3), 379-400.

Harley, B., & Hart, D. (2002). Age, aptitude, and second language learning on a bilingual exchange. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 301-330). Amsterdam: John Benjamins.

Harrington, M. (2004). Commentary: Input processing as a theory of processing input. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 79-92). Mahwah, NJ: Lawrence Erlbaum.

Hedgcock, J. (1993). Well-formed vs. ill-formed strings in L2 metalingual tasks: Specifying features of grammaticality judgements. *Second Language Research, 9*, 1-21.

Heilenman, L. (1983). The use of a cloze procedure in foreign language placement. *The Modern Language Journal, 67*, 121-126.

Hinkel, E. (2002). Why English passive is difficult to teach (and learn). In E. Hinkel & S. Fotos (Eds.), *New perspectives on grammar teaching in second language classrooms* (pp. 233-259). Mahwah, NJ: Lawrence Erlbaum.

Hinkel, E. (2004). Tense, aspect and the passive voice in L1 and L2 academic texts. *Language Teaching Research, 8*(1), 5-29.

Housen, A., Pierrard, M., & Van Daele, S. (2005). Structure complexity and the efficacy of explicit grammar instruction. In A. Housen & M. Pierrard (Eds.), *Investigations in instructed second language acquisition* (pp. 235-269). Brusells, Belguim: Mouton De Gruyter.

Hu, G. (2002). Psychological constraints on the utility of metalinguistic knowledge in second language production. *Studies in Second Language Acquisition, 24*(3), 347-386.

Hulstijn, J. (1995). Not all grammar rules are equal: Giving grammar instruction its proper place in foreign language teaching In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 359-386). Honolulu, HI: University of Hawai'i Press.

Hulstijn, J. (2002). Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research, 18*, 193-223.

Hulstijn, J. (2005). Theoretical and empirical issues in the study of implicit and explicit second language learning: Introduction. *Studies in Second Language Acquisition, 27*(2), 129-140.

Hulstijn, J., & de Graaff, R. (1994). Under what conditions does explicit knowledge of a second language facilitate the acquisition of implicit knowledge? A research proposal. *AILA Review, 11*, 97-112.

Hulstijn, J., & Marchena, E. (1989). Avoidance: Grammatical or semantic causes. *Studies in Second Language Acquisition, 11*, 242-255.

Hwang, M. (1979). *A semantic and syntactic analysis of 'If conditionals".* Unpublished doctoral dissertation, University of California, Los Angeles, California.

Izquierdo, J., & Collins, L. (2008). The facilitative role of L1 influence in tense-aspect marking: A comparison of Hispanophone and Anglophone learners of French. *Modern Language Journal, 92*, 350-368.

Izumi, S., & Lakshmanan, U. (1998). Learnability, negative evidence and the L2 acquisition of the English passive. *Second Language Research, 14*(1), 62-101.

Izumi, Y., & Izumi, S. (2004). Investigating the effects of oral output on the learning of relative clauses in English: Issues in the psycholinguist requirements for effective output tasks. *The Canadian Modern Language Review, 60*(5), 587-609.

Keenan, E., & Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry, 8*(1), 63-99.

Kellerman, E., & Sharwood-Smith, M. (Eds.). (1986). *Cross-linguistic influence in second language acquisition*. London, England: Pergamon.

Kleinmann, H. (1978). The strategy of avoidance in adult second language acquisition. In W. Ritchie (Ed.), *Second language acquisition research: Issues and implications* (pp. 157-174). New York, NY: Academic Press.

Krashen, S. (1981). *Second language acquisition and second language learning* London, England: Pergamon.

Krashen, S. (1982). *Principles and practice in second language acquisition*. London, England: Pergamon.

Krashen, S. (1985). *The input hypothesis: Issues and implications*. London, England: Longman.

Krashen, S. (1994). The input hypothesis and its rivals. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 45-77). London, England: Academic Press.

Krashen, S., & Terrell, T. (1983). *The natural approach*. New York, NY: Pergamon.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor, MI: University of Michigan.

Laesch, K., & van Kleeck, A. (1987). The cloze test as an alternative measure of language proficiency of children considered for exit from bilingual education programs. *Language Learning, 37*, 171-189.

Langacker, R. (1987). *Foundations of cognitive grammar, Vol. I.* Stanford, CA: Stanford University Press.

Larsen-Freeman, D. (1995). On the teaching and learning of grammar: Challenging the myths. In F. Eckman, D. Highland, P. Lee, J. Mileham & R. Rukowski Weber (Eds.), *Second language acquisition theory and pedagogy* (pp. 131-168). Mahwah, NJ: Lawrence Erlbaum.

Larsen-Freeman, D. (2003a). The grammar of choice. In E. Hinkel & S. Fotos (Eds.), *New perspectives on grammar teaching*. Mahwah, NJ: Lawrence Erlbaum.

Larsen-Freeman, D. (2003b). *Teaching language: From grammar to grammaring*. Boston, MA: Thomson & Heinle.

Leow, R. (1997a). Attention, awareness, and foreign language behavior. *Language Learning, 47*, 467-505.

Leow, R. (1997b). The effects of input enhancement and text length on adult readers' comprehension and intake in second language acquisition. *Applied Language Learning, 8*, 151-182.

Leow, R. (2000). A study of the role of awareness in foreign language behavior: Aware versus unaware learners. *Studies in Second Language Acquisition, 22*, 557-582.

Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge, MA: Massachusetts Institute of Technology Press.

Li, C., & Thompson, S. (1981). *Mandarin Chinese*. Berkeley, CA: University of California Press.

Lightbown, P. M., & Spada, N. (2000). Do they know what they're doing? L2 learners' awareness of L1 influence. *Language Awareness, 9*(4), 198-216.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Thousand Oaks, CA: Sage.

Loewen, S. (2009). Grammaticality judgment tests and the measurement of implicit and explicit L2 knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 94-112). Tonawanda, NY: Multilingual Matters.

Loewen, S., Li, S., Fei, F., Thompson, A., Nakatsukasa, K., Ahn, S., et al. (2009). Second language learners' beliefs about grammar instruction and error correction. *Modern Language Journal, 93*, 93-104.

Long, M. (1983). Does second language instruction make a difference? A review of research. *TESOL Quarterly, 17*(3), 359-382.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego, CA: Academic Press.

Long, M. (2007). Recasts in SLA: The story so far. In M. Long (Ed.), *Problems in SLA* (pp. 75-116). Mahwah, NJ: Lawrence Erlbaum.

Loschky, L., & Bley-Vroman, R. (1993). Grammar and task-based methodology. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 123-167). Clevedon, England: Multilingual Matters.

Luk, Z. P., & Shirai, Y. (2009). Is the acquisition order of grammatical morphemes impervious to L1 knowledge? Evidence from the acquisition of plural -s, articles, and possessive 's. *Language Learning, 59*(4), 721-754.

Lyster, R. (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition, 26*, 399-432.

Lyster, R. (2007). Complementary roles for input and output enhancement in form-focused instruction. In C. Gascoigne (Ed.), *Assessing the impact of input enhancement in second language education: Evolution in theory, research and practice* (pp. 129-151). Stillwater, OK: New Forums.

Lyster, R., & Izquierdo, J. (2009). Prompts versus recasts in dyadic interaction. *Language Learning 59*(2), 453-498.

Macaro, E., & Masterman, L. (2006). Does intensive explicit grammar instruction make all the difference? *Language Teaching Research, 10*(3), 297-327.

Mackey, A. (1999). Input, interaction and second language development: An empirical study of question formation in ESL. *Studies in Second Language Acquisition, 21*, 557-587.

Mackey, A., & Philp, J. (1998). Recasts, interaction, and interlanguage development: Are responses red herrings? *Modern Language Journal, 82*, 338-356.

Mackey, A., Philp, J., Egi, T., Fujii, A., & Tatsumi, T. (2002). Individual differences in working memory, noticing of interactional feedback and L2 development. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 181-209). Amsterdam: John Benjamins.

Master, P. (1991). Active verbs with inanimate subjects in scientific prose. *English for Specific Purposes, 10*, 15-33.

Master, P. (1994). The effect of systematic instruction on learning the English article system. In T. Odlin (Ed.), *Perspectives on pedagogical grammar* (pp. 229-252). New York, NY: Cambridge University Press.

McDonough, K. (2007). Interactional feedback and the emergence of simple past activity verbs in L2 English. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 323-338). Oxford, England: Oxford University Press.

McEnery, T., & Xiao, R. (2005). Passive constructions in English and Chinese: A corpus-based contrastive study. Retrieved January 20, 2010, from http://eprints.lancs.ac.uk/63/

McNamara, T. (1996). *Measuring second language performance*. London, England: Longman.

Meisel, J., Clahsen, H., & Pienemann, M. (1981). On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition, 3*, 109-135.

Mellow, J. D. (2006). The emergence of second language syntax: A case study of the acquisition of relative clauses. *Applied Linguistics, 27*(4), 645-670.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13-103). New York, NY: Macmillan.

Messick, S. (1996). Validity and washback in language testing. *Language Testing, 13*, 241-256.

Norris, J., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning, 50*(3), 417-528.

Norris, J., & Ortega, L. (2003). Defining and measuring SLA. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 256-310). Oxford, England: Blackwell.

Odlin, T. (1989). *Language transfer*. Cambridge, England: Cambridge University Press.

Odlin, T. (2003). Cross-linguistic influence In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 436-486). New York, NY: Blackwell.

Oller, J. W. J., & Conrad, C. A. (1971). The cloze technique and ESL proficiency. *Language Learning, 21*, 183-195.

Ortega, L. (2009). *Understanding second language acquisition*. London, England: Hodder Education.

Paradis, M. (1994). Neurolinguistic aspects of implicit and explicit memory: Implications for bilingualism and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 393-419). San Diego, CA: Academic Press.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.

Paradis, M. (2009). *Declarative and procedural determinants of second languages*. Amsterdam: John Benjamins.

Pavesi, M. (1986). Markedness, discoursal modes, and relative clause formation in a formal and an informal context. *Studies in Second Language Acquisition, 8*, 38-55.

Pienemann, M. (1989). "Is language teachable?" Psycholinguistic experiments and hypotheses. *Applied Linguistics, 10*, 52-79.

Pienemann, M. (1998). *Language processing and second language development: Processability Theory*. Amsterdam: John Benjamins.

Pienemann, M. (2003). Language processing capacity. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 679-714). Oxford, England: Blackwell.

Pienemann, M. (2005). An introduction to processability theory. In M. Pienemann (Ed.), *Cross-linguistic aspects of processability theory* (pp. 1-60). Amsterdam: Benjamins.

Pienemann, M., Johnson, M., & Brindley, G. (1988). Constructing an acquisition-based procedure for assessing second language acquisition. *Studies in Second Language Acquisition, 10*, 217-243.

Ranta, L. (2002). The role of learners' language analytic ability in the communicative classroom. In P. Robinson (Ed.), *Individual differences and instructed second language acquisition*. Amsterdam: John Benjamins.

Reber, A. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219-235.

Robinson, P. (1995a). Aptitude, awareness, and the fundamental similarity of implicit and explicit second language learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 303-358). Hawaii, HI: University of Hawaii Second Language Teaching and Curriculum Center.

Robinson, P. (1995b). Attention, memory, and the noticing hypothesis. *Language Learning, 45*(2), 283-331.

Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition, 18*, 27-68.

Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 285-317). Cambridge, England: Cambridge University Press.

Robinson, P. (2002a). *Individual differences and instructed language learning*. Amsterdam: John Benjamins.

Robinson, P. (2002b). Learning conditions, aptitude complexities and SLA: A framework for research and pedagogy. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 113-133). Amsterdam: John Benjamins.

Robinson, P. (2003). Attention and memory during SLA. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 631-678). Malden, MA: Blackwell.

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics, 25*, 46-73.

Roehr, K. (2006). Metalinguistic knowledge in L2 task performance: A verbal protocol analysis. *Language Awareness, 15*(3), 180-198.

Roehr, K. (2007). Metalinguistic knowledge and language ability in university-level L2 learners. (Publication no. 10.1093/applin//amm037). Retrieved May, 30, 2008, Oxford University Press.

Roehr, K. (2008a). Linguistic and metalinguistic categories in second language learning. *Cognitive Linguistics, 19*(1), 67-106.

Roehr, K. (2008b). Metalinguistic knowledge and language ability in university-level L2 learners. *Applied Linguistics, 29*(2), 173-199.

Roehr, K., & Ganem-Gutierrez, G. (2009). Metalinguistic knowledge: A stepping stone towards L2 proficiency. In A. Benati (Ed.), *Issues in second language proficiency* (pp. 79-94). New York, NY: Continuum.

Rosa, E., & Leow, R. (2004). Awareness, different learning conditions, and L2 development. *Applied Psycholinguistics, 25*, 269-292.

Rosa, E., & O'Neill, M. (1999). Explicitness, intake, and the issue of awareness: Another piece of the puzzle. *Studies in Second Language Acquisition, 21*, 511-533.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133-164). Amsterdam: John Benjamins.

Sawyer, M., & Ranta, L. (2001). Aptitude, individual differences, and instruction design. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 319-353). Cambridge, England: Cambridge University Press.

Schachter, J. (1974). An error in error analysis. *Language Learning, 27*, 205-214.

Schachter, J., & Yip, V. (1990). Grammaticality judgments: Why does anyone object to subject extraction? *Studies in Second Language Acquisition, 12*, 379-392.

Scheffler, P. (2009). Rule difficulty and the usefulness of instruction. *ELT Journal, 63*, 5-12.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*, 17-46.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, 11-26.

Schmidt, R. (1995). Consciousness and foreign language learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1-63). Honolulu, HI: University of Hawaii Press.

Schmidt, R. (2001). Attention In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge, England: Cambridge University Press.

Schmierer, R. (1979). Linguistic complexity and the teaching of English structure. *Language Learning, 29*(2), 243-254.

Schulz, R. A. (1996). Focus on form in the foreign language classroom: Students' and teachers' views on error correction and the role of grammar. *Foreign Language Annals, 29*(3), 340-353.

Schulz, R. A. (2001). Cultural differences in student and teacher perceptions concerning the role of grammar instruction and corrective feedback: USA-Colombia. *Modern Language Journal, 85*(2), 244-258.

Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behavior. *Studies in Second Language Acquisition, 15*, 147-163.

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in second language learning. In J. F. Kroll & A. M. B. DeGroot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371-388). Oxford: Oxford University Press.

Selinker, L., & Lakshamanan, U. (1992). Language transfer and fossilization. In S. Gass & L. Selinker (Eds.), *Language transfer in language learning* (2nd ed., pp. 196-215). Philadelphia, PA: John Benjamins.

Sharwood-Smith, M. (1981). Consciousness-raising and the second language learner. *Applied Linguistics, 2*, 159-168.

Sheen, Y. (2007a). The effect of focused written corrective feedback and language aptitude on ESL learners' acquisition of articles. *TESOL Quarterly, 41*, 255-283.

Sheen, Y. (2007b). The effects of corrective feedback, language aptitude and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction and second language acquisition* (pp. 301-332). Oxford, England: Oxford University Press.

Shirai, Y. (2004). A multi-factor account for form-meaning connections in the acquisition of tense-aspect morphology. In B. VanPatten, J. Williams, S. Rott & M. Overstreet (Eds.), *Form-meaning connections in second language acquisition* (pp. 97-122). Mahwah, NJ: Lawrence Erlbaum.

Shirai, Y., & Ozeki, H. (2007). Special issue: The L2 acquisition of relative clauses in East Asia languages. *Studies in Second Language Acquisition, 29*, 155-374.

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford, England: Oxford University Press.

Skehan, P. (2001). Tasks and language performance assessment. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks, second language learning, teaching and testing* (pp. 167-185). London, England: Longman.

Skehan, P. (2002). Theorising and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69-93). Amsterdam: John Benjamins

Snow, R. E. (1987). Aptitude complexes. In R. E. Snow & M. J. Farr (Eds.), *Aptitude, learning and instruction* (pp. 13-59). Hillsdale, NJ: Lawrence Erbaum.

Snow, R. E. (1994). Abilities in academic tasks. In R. Sternberg & R. K. Wagner (Eds.), *Mind in context: Interactionist perspective on human intelligence* (pp. 3-37). New York, NY: Cambridge University Press.

Sorace, A. (1985). Metalinguistic knowledge and language use in acquisition-poor environments. *Applied Linguistics, 6*(3), 239-254.

Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching, 30*, 73-87.

Spada, N., & Lightbown, P. M. (1993). Instruction and the development of questions in the L2 classroom. *Studies in Second Language Acquisition, 15*, 205-221.

Spada, N., & Lightbown, P. M. (1999). Instruction, L1 influence and developmental readiness in second language acquisition. *Modern Language Journal, 83*, 1-22.

Spada, N., Lightbown, P. M., & White, J. (2005). The importance of meaning in explicit form-focused instruction. In A. Housen & M. Pierrard (Eds.), *Current issues in instructed second language learning* (pp. 199-234). Brussels, Belgium: Mouton De Gruyter.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60*(2), 263-308.

Spada, N., Jessop, L., Suzuki, W., Tomita, Y., & Valeo, A. (In preparation). *Isolated and integrated form-focused instruction: Equally effective for different types of L2 knowledge and ability?* Unpublished manuscript, University of Toronto.

Stockwell, R. P., Bowen, J. D., & Martin, J. W. (1965). *The grammatical structures of English and Spanish*. Chicago, IL: University of Chicago Press.

Swain, M. (2005). The output hypothesis: Theory and research In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471-481). Mahwah, NJ: Lawrence Erlbaum.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th. ed.). New York, NY: Harper & Row.

Thomas, M. (1994). Assessment of L2 proficiency in second language acquisition research. *Language Learning, 44*(2), 307-336.

Thomas, M. (2006). Research synthesis and historiography: The case of assessment of second language proficiency. In J. M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 279-298). Amsterdam: John Benjamins.

Tomlin, R., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*, 183-203.

Trofimovich, P., Ammar, A., & Gatbonton, B. (2007). How effective are recasts? The role of attention, memory, and analytical ability. In A. Mackey (Ed.), *Conversational interaction in second language acquisition* (pp. 171-195). Oxford, England: Oxford University Press.

Valeo, A. (2010). *The integration and language and content: Form-focused instruction in a content-based language program.* Unpublished doctoral dissertation, University of Toronto, Toronto.

Van Baalen, T. (1983). Giving learners rules: A study into the effect of grammatical instruction with varying degrees of explicitness. *Interlanguage Studies Bulletin Utrecht, 7*, 71-100.

VanPatten, B. (1996). *Input processing and grammar instruction*. New York, NY: Ablex.

VanPatten, B. (2002). Processing instruction: An update. *Language Learning, 52*(4), 755-803.

VanPatten, B. (2007). Some thoughts on the future of research on input enhancement. In C. Gascoigne (Ed.), *Assessing the impact of input enhancement in second language education: Evolution in theory, research and practice* (pp. 169-189). Stillwater, OK: New Forum.

VanPatten, B. (Ed.). (2004). *Processing instruction: Theory, research and commentary*. Mahwah, NJ: Lawrence Erlbaum.

Wesche, M. (1981). Language aptitude measures in streaming, matching students with methods, and diagnosis of learning problems. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 119-154). Rowley, MA: Newbury House.

White, J. (1998). Getting the learners' attention: A typographical input enhancement study. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom language acquisition* (pp. 85-113). Cambridge, England: Cambridge University Press.

White, J., Munoz, C., & Collins, L. (2007). The his/her challenge: Making progress in a 'regular' L2 programme. *Language Awareness, 16*(4), 278-299.

White, J., & Ranta, L. (2002). Examining the interface between metalinguistic task performance and oral production in a second language. *Language Awareness, 11*(4), 259-290.

White, L. (1991). Adverb placement in second language acquisition: Some effects of positive and negative evidence in the classroom. *Second Language Research, 7*, 133-161.

Williams, J. (2005). Form-focused instruction. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 671-692). New Jersey, NJ: Lawrence Erlbaum.

Williams, J., & Evans, J. (1998). What kind of focus and on which forms? In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 139-155). Cambridge, England: Cambridge University Press.

Williams, J. N. (1999). Memory, attention, and inductive learning. *Studies in Second Language Acquisition, 21*, 1-48.

Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition, 27*, 269-304.

Wong, W. (2004). Processing instruction in French: The roles of explicit information and structured input. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 187-206). Mahwah, NJ: Lawrence Erlbaum.

Wu, C. (1994). *"If triangles were circles, ..." - a study of counterfactuals in Chinese and in English*. Taipei, Taiwan: Crane.

Yip, V. (1995). *Interlanguage and learnability: From Chinese to English*. Amsterdam: John Benjamins.

Yuan, F., & Ellis, R. (2003). The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics, 24*(1), 1-27.

Zobl, H. (1985). Grammar in search of input and intake. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 329-344). Rowley, MA: Newbury House.

Appendix A

Questionnaire

Section 1: Background Information.

Please put a check mark (√) on the line that is most appropriate for your situation and provide written responses (in Chinese or English) where necessary.

1. Sex ☐ Male ☐ Female

2. First Language ☐ Taiwanese ☐ Mandarin Chinese ☐ Others _____

3. I started learning English when I was a ___
   ☐ kindergartener ☐ first grader ☐ second grader ☐ third grader
   ☐ fourth grader ☐ fifth grader ☐ sixth grader ☐ seventh grader

4. Age _____ years old

5. Have you ever lived in an English-speaking country?
   ☐ No ☐ Yes. (For how long? _____ Years _____ Months)

6. I am currently a _____ student.
   ☐ first-year ☐ second-year ☐ third-year ☐ fourth-year
   ☐ graduate

7. I am a/an ____.
   ☐ English-major ☐ non-English-major (Department _____)

8. How often have you been taught grammar rules in English classes?

At elementary school	☐ Very often	☐ Often	☐ Sometimes	☐ Seldom
At Junior high	☐ Very often	☐ Often	☐ Sometimes	☐ Seldom
At Senior high	☐ Very often	☐ Often	☐ Sometimes	☐ Seldom
At University	☐ Very often	☐ Often	☐ Sometimes	☐ Seldom

9. How would you assess your general English grammar knowledge?
   ☐ Very good ☐ Good ☐ Somewhat poor ☐ Poor

10. What is your score for the English subject in the Joint College Entrance Examination?
    My score was _____.

11. Have you taken any English proficiency tests? (multiple choice) (Please indicate the score that you got for the test(s) you have taken, if available.)

☐ No, never.

☐ Yes.

☐ I have taken GEPT, but I did not pass.

☐ I have taken GEPT, and I have passed ____.

☐ the listening and reading parts of the beginning level of GEPT (Score _____)

☐ the speaking and writing parts of the beginning level of GEPT (Score _____)

☐ the listening and reading parts of the intermediate level of GEPT (Score _____)

☐ the speaking and writing parts of the intermediate level of GEPT (Score _____)

☐ the listening and reading parts of the high-intermediate level of GEPT (Score _____)

☐ the speaking and writing parts of the high-intermediate level of GEPT (Score _____)

☐ the listening and reading parts of the advanced level of GEPT (Score _____)

☐ the speaking and writing parts of the advanced level of GEPT (Score _____)

☐ I have taken TOFEL CBT, and I have got _____.

☐ I have taken TOEFL iBT, and I have got _____.

☐ I have taken TOEFL test, and I have got _____.

☐ I have taken IELTS, and I have got _____.

Others (Please specify_____)

12. Do you have additional chances to be exposed to English or to use English out of English classes?

☐ No.

☐ Yes.

☐ I have traveled to an English-speaking country

☐ I communicate online in English (in chat rooms or with an email partner)

☐ I often talk with my friends or family in English

☐ others (please specify _____)

Section 2: Learners' Perceptions of Grammatical Difficulty.

The following is a list of English grammatical features. Please indicate whether a particular grammatical feature has been more or less difficult for you to learn by circling **only one** number between 1 and 6. And please be sure to respond to all the grammatical features.

Note:

"Not at all difficult" indicates that you have learned the structure quickly after a short explanation and practice.

"Extremely difficult" indicates that you never expect to learn the structure fully, even with sufficient explanation and practice.

		Not at all difficult ⟷ Extremely difficult					
1	Present perfect • I have finished the job. • I have washed my father's car.	1	2	3	4	5	6
2	Simple past -ed • She looked very happy yesterday. • I talked to my teacher two days ago.	1	2	3	4	5	6
3	Negation • He is not a teacher. • He does not like swimming.	1	2	3	4	5	6
4	Modal verbs (may/might, can/could) • He may come tomorrow. • He can speak English very well.	1	2	3	4	5	6
5	Countable and uncountable nouns • This is a book. (countable noun) • I like rice. (uncountable noun)	1	2	3	4	5	6

6	Passives • The package was delivered to Taiwan. • Today many products are made of plastic.	1	2	3	4	5	6
7	Articles (a, an, the) • It is a book. • He is in the classroom.	1	2	3	4	5	6
8	Unreal conditionals • If she had money, she would travel. • If I were a bird, I would fly high.	1	2	3	4	5	6
9	Embedded questions • He asked where John lived. • He does not know what his name is.	1	2	3	4	5	6
10	Third person -s • He lives in Taipei. • The woman works at the hospital.	1	2	3	4	5	6
11	Clauses • The cake that you bought yesterday is delicious. • It is true that the earth is round.	1	2	3	4	5	6
12	Present Progressive • John is playing a video game. • Mary is dancing.	1	2	3	4	5	6
13	Prepositions • The book is on the table. • I will meet you at six.	1	2	3	4	5	6
14	Adjective comparatives • He is taller than Mary. • He is more handsome than John.	1	2	3	4	5	6

15	Infinitive • To become a rock star is not easy. • We agreed to play in the band.	1	2	3	4	5	6
16	Real conditionals • If you stay up late, you will feel tired tomorrow. • If it rains tomorrow, I will stay home.	1	2	3	4	5	6
17	Wh-questions • What is your name? • Where do you live?	1	2	3	4	5	6
18	Participial construction • Spoken in many countries, English is a universal language. • Watching TV, she forgot everything around her.	1	2	3	4	5	6
19	Question tags • It's beautiful, isn't it? • You don't like fish, do you?	1	2	3	4	5	6
20	Past progressive • I was studying at 8:00 last night. • He was reading a book in bed.	1	2	3	4	5	6

Section 3. Please indicate whether the selected target grammar structures have been more or less difficult for you to learn, and then please explain why.

1. I think that "third person -s" has been ____ for me to learn.

___ (1) not at all difficult ___ (2) a little bit difficult

___ (3) difficult ___ (4) very difficult

(Sample sentences for "third person -s" → *He lives in Taipei. The woman works at the hospital.*).

The reason for my choice is that

2. I think that "passives" have been ____ for me to learn.

___ (1) not at all difficult ___ (2) a little bit difficult

___ (3) difficult ___ (4) very difficult

(Sample sentences for "passives" → *The package was delivered to Taiwan. Today many products are made of plastic.*)

The reason for my choice is that

(Please turn to next page)

3. I think that “articles” have been ____ for me to learn.

___ (1) not at all difficult ___ (2) a little bit difficult
___ (3) difficult ___ (4) very difficult
(Sample sentences for “articles” → *It is <u>a</u> book. He is in <u>the</u> classroom*.)

The reason for my choice is that

4. I think that “present perfect” has been ____ for me to learn.

___ (1) not at all difficult ___ (2) a little bit difficult
___ (3) difficult ___ (4) very difficult
(Sample sentences for “present perfect” → *I <u>have finished</u> the job. I <u>have washed</u> my father’s car.*
The reason for my choice is that

Thank you very much for participating in this questionnaire survey.

If you would like to participate in a follow-up interview and an oral language production test, please leave your contact information below.

Cell phone number: _____(optional) E-mail: _____

Appendix B

Proficiency Test

(Reprinted in this thesis with permission)

Directions: Complete the blanks in the following text. You have a maximum of 45 minutes for this test.

I loved the smallest lion cub best. She was the weakest, but she was the bravest of the three. I called her Elsa. If the cubs had stayed in the _(l)_____, Elsa would be dead. A lioness _(2)_____ has four cubs. One dies soon _(3)_____ birth, and another one is often _(4)_____ weak to live long. It is _(5)_____ this reason that you usually see _(6)_____ two cubs with a lioness. Their _(7)_____ looks after them well until they _(8)_____ two years old. For the first _(9)_____ she brings their food. She eats _(10)_____ herself first, then she vomits it _(11)_____ again. In this way, the cubs _(12)_____ manage the food.

Lions often live _(13)_____ in a group called a "pride". _(14)_____ full-grown lions of the pride _(15)_____ all of the hunting. During their _(16)_____ year the cubs hunt with the _(17)_____, but they are unable to kill _(18)_____ their own. They eat only what _(19)_____ big lions leave. Often very little _(20)_____ for them, so they are usually _(21)_____ a very bad condition at this _(22)_____.

Sometimes the hunger is too much _(23)_____ the cubs. They will try to take the _(24)_____ before the big lions have finished _(25)_____. The big lions will kill them _(26)_____ this. Sometimes the hungry cubs will _(27)_____ the pride in small groups. And _(28)_____ they do not yet know how _(29)_____ kill, they often have trouble. Nature's _(30)_____ is hard, and lions have to _(31)_____ the hard way from the beginning.

_(32)_____ the wild, Elsa would be the throw-out _(33)_____ the pride. At five months old, _(34)_____ with us, she was in fine _(35)_____. But she was still weaker than the _(36)_____ two cubs. They fought over every _(37)_____, and poor Elsa did not get _(38)_____ to eat. I kept bits of _(39)_____ for her and held her on

_(40)_____ lap for all her meals. I _(41)_____ with her while I fed her, _(42)_____ she loved this. During these hours, _(43)_____ deep feeling grew between us.

We could _(44)_____ keep three fast-growing lions forever. _(45)_____, we decided that the two big _(46)_____ had to go. I chose to _(47)_____ Elsa, and everyone in the house _(48)_____. Perhaps they were thinking, "If there _(49)_____ be a lion in the house, _(50)_____ should be as small as possible". We sent the two biggest cubs away to Europe.

Appendix C
Letter of Inquiry for Students

December X, 2008

Dear Student,

I am a Ph.D. student in the Second Language Education Program at the Ontario Institute for Studies in Education at the University of Toronto (OISE/UT), and I am conducting a research project on Chinese students' English grammar learning. I am looking for students who are willing to participate in this study. The main goal of the research is to explore (1) the grammatical features university-level Chinese learners of English as a foreign language (EFL) perceive as more or less difficult to learn, (2) why learners consider some features to be more difficult than others, and (3) whether learners' perceptions of grammatical difficulty correlate with their second language (L2) proficiency and their actual L2 performance. The findings of the study will positively contribute to theory and practice in second or foreign language instruction.

I hope that your participation in the study will help you to think about your English grammar learning. At the same time, the information that you provide will help researchers and teachers who teach English as a second or foreign language increase their knowledge of how Chinese students learn English grammar.

Participation in this study is completely voluntary. If you agree to participate, you may withdraw from the study at any time, for any reason. Also, there would be no negative consequences if you decide to have any of data collected pertaining to you removed from the study.

The study involves a questionnaire survey, a cloze test, an interview, an oral language production test, and a written grammar test. For more details about the data collection procedures of the study, please refer to the enclosed "Description of the Research Project."

Should you agree to participate in the study or only part of the study, you will be asked to do some of the following:

– to sign a consent form affirming that you fully understand the purpose and the procedures of the study, and to keep a copy for your own references.

– to respond to a written questionnaire asking your ideas about what grammatical features are more or less difficult to learn. The questionnaire will also ask you to provide biographical data and information about yourself and your previous English learning experiences. At the end of the questionnaire, you will be asked whether you are willing to participate in a follow-up interview and an oral language production test. The administration of the questionnaire will be in one of your required English classes. The questionnaire is estimated to take approximately 40 minutes to complete.
– to participate in a cloze test in one of your required English classes after completion of the questionnaire. The class visit for the cloze test will be arranged in consultation with your English instructor. The cloze test consists of 50 blanks that require you to fill in the blanks with words that you think are appropriate according to the context. The cloze test is estimated to take a maximum of 45 minutes to complete.

– to participate in an individual interview, taking approximately one hour, to talk about your previous grammar learning experiences and your opinions about the degree of difficulty of selected features of English grammar. During the interview, you will be asked to do an activity that requires you to rank 10 selected grammatical features according to their difficulty level and then to explain your ranking. In addition, you will also be asked to do two short cloze activities, which are intended to help you recall your previous experiences of learning specific grammar features. The interview session will be conducted at your convenience, outside your regular class hours. The interview will be audio-recorded for later transcription and reference, and will be conducted in Mandarin Chinese. Note: Approximately 25 to 30 students will be invited to participate in the interview sessions. The interview will be conducted two to three weeks after the questionnaire survey. Upon completion of the interview, in return for their time spent on the interview, participants will receive a gift coupon with a value equivalent to 100 NT dollars (approximately 3.3 Canadian dollars).

–to participate in an oral language production test that will take approximately 15 to 20 minutes. For this part of research, I will individually meet with you, at your convenience, outside of your regular class hours. During the test, you will be presented with 2 sets of 8 to 10 pictures. You will be asked to describe the pictures in English, using the cues provided by the researcher and the pictures themselves. The whole process will be audio-recorded. Note: Approximately 25 to 30 students that are interviewed will be invited to participate in the oral test. The test will be carried out two to three months after the completion of the questionnaire survey. Upon completion of the task, in return for their time spent on the task, participants will receive a gift coupon with a value equivalent to 100 NT dollars (approximately 3.3 Canadian dollars).

– to participate in a written grammar test in one of your required English classes. The test consists of 40 ungrammatical sentences, which you will be asked to correct, and to explain your corrections. The grammar test is estimated to take a maximum of 60 minutes. Note: Approximately 3 or 4 English classes that complete the questionnaire will be invited to take the test. The grammar test will be administered some time after the completion of the oral picture-description task, and they will be administered at the convenience of your instructor.

- to allow Ms. Shiu to use the data that you provide in any of her publications or presentations associated with this study. The data that you provide will be reported in a manner that your identity will not be discerned.

– to keep your (and your classmates') participation information confidential.

I would like to emphasize again that your participation in this study is completely voluntary. Your decision either to participate or not to participate will not affect your grade. You are not obliged to answer any questions you find objectionable, and you are free to withdraw from the study at any point, and to request the removal of all or part of the data pertaining to you.

All the information you provide will be kept fully confidential. In order to conceal your identity, I will use a code number and/or a pseudonym to store and analyze the data you provide, and in reporting the results of the study. Only my thesis supervisor (Professor Nina Spada) and I will have access to your personal information. No information including names and other indications of your identity will be shared with anybody in your department, including your classmates, your teachers and the administrators.
I will provide you with a summary of the findings from the study after the thesis is completed if you wish. You can indicate your wish to receive the summary on the consent form.

If you choose to participate, please sign the Informed Consent Form, a copy of which will be provided to you. If you have any questions, please do not hesitate to contact me by telephone (002-1-416-312-XXXX, when I am in Canada or 0963-182-XXX, when I am in Taiwan) or by email (lishiu@oise.utoronto.ca). You can also reach my supervisor, Dr. Nina Spada, by email (nspada@oise.utoronto.ca). Should you have any questions about participant rights, please feel free to contact the Office of Research Ethics at University of Toronto by email (ethics.review@utoronto.ca) or by telephone (002-1-416-946-3273).

Yours sincerely,
Li-Ju Shiu

PhD Candidate, Second Language Education
Department of Curriculum, Teaching and Learning
OISE/University of Toronto
Canada

Appendix D

Consent Form for the Students

(To be placed on a departmental letterhead)

CONSENT FORM

I have read Li-Ju Shiu's Letter of Inquiry describing her research project entitled *Investigating Grammatical Difficulty in Relation to Proficiency Level and Different Knowledge Type in Second Language Learning,* and I fully understand the purpose and the procedures of the study (including the measures to be taken to maintain confidentiality).

In the following, please check (√) the item(s) that you agree to participate in and that you are aware of.

For this research project, I agree:

- ☐ to respond to a questionnaire survey asking for: (1) information about me and my previous English learning experience; (2) my ideas about what grammatical features are more or less difficult to learn.; and (3) my willingness to participate in a follow-up individual interview and an oral language production test. The questionnaire will be administered in one of my required English classes. The questionnaire is estimated to take approximately 40 minutes to complete.
- ☐ to participate in a cloze test that requires me to fill in the blanks with words that I think are appropriate according to the text context. The cloze test will be administered in one of my English classes. The cloze test is estimated to take a maximum of 45 minutes to complete.
- ☐ to be interviewed for approximately one hour about my previous grammar learning experience and my ideas about grammar learning. During the interview, I will also be asked to do an activity and two short cloze activities that are intended to help me recall learning specific grammar features.
- ☐ to have my interview session digitally audio-recorded.

☐ to participate in an oral language production test that requires me to use English to describe pictures presented to me. I will individually meet with Ms. Shiu, at my convenience, for the oral language production test. The oral test will take approximately 20 minutes to complete.

☐ to have my oral language production test digitally audio-recorded.

☐ to take a written grammar test that consists of 40 ungrammatical sentences in which I will correct the errors and explain my corrections. The grammar test will be administered in one of my English classes. The grammar test will take a maximum of 60 minutes to complete.

☐ to allow Ms. Shiu to use the data that I provide in any of her publications or presentations associated with this study. The data that I provide will be reported in a manner such that my identity will not be disclosed.

☐ to keep my (and my classmates') participation information confidential.

And I am also aware that

☐ even if I agree to participate in the study, I have the right to withdraw from the study or part of the study, without any consequence, at any point.

☐ if I participate in the individual interview, I can refuse to answer any of the interview questions for any reason.

☐ if I participate in the interview, during the interview, I can request to stop the recording at any point for any reason.

☐ if I participate in the oral language production test, during the test, I can request to stop the recording at any point for any reason.

☐ if I take part in the individual interview I will receive 100 New Taiwan dollars in return for my time.

☐ if I take part in the oral language production test I will receive 100 New Taiwan dollars in return for my time.

☐ if I have any questions about participant rights, I can contact the Office of Research Ethics at University of Toronto by email (ethics.review@utoronto.ca) or by telephone (002-1-416-3273).

After Ms. Shiu completes her study, (please check (√) the following)

☐ I do not wish to receive a summary of the research report.

☐ I wish to receive a summary of the research report.

(Please email the summary to _____ OR mail it to

_____)

Signature: _____ Date: _____

Appendix E

Interview Protocol

Before the interview starts, the interviewee will be informed that

1. This interview contains questions regarding your attitude toward grammar learning, preferences for grammar learning, and prior grammar learning experiences.
2. You can refuse to answer any of the interview questions during the interview if you are not comfortable responding to any particular question(s).
3. You can ask any questions related to this research.
4. This interview will be audio-recorded and transcribed later for the data analysis. You can request to stop the recorder anytime you feel uncomfortable being recorded during the interview.
5. You can be assured that the interview data will be kept confidential and used only for the research purpose of this study.

Questions about Grammar Learning

1. How have you been taught grammar in high school?
2. Do you like studying grammar?
3. Does learning grammar help your reading, writing, listening or speaking?
4. Do you think that learning grammar is important for mastering English? Why or why not?

Appendix F

Cloze Activities

Cloze activity A

Statue of Liberty

Read the passage and fill in the blanks with the correct form of the verb in the brackets.

The Statue of Liberty is 46 metres high. The statue represents the goddess of liberty. She holds a torch in her right hand and a tablet in her left hand with the date of the Declaration of Independence on it. The statue _____ (give) to the United States by France. It was a present on the 100th anniversary of the United States. The Statue of Liberty _____ (design) by Frederic Auguste Bartholdi and it _____ (complete) in France in July 1884. In 350 pieces, the statue _____ (ship) to New York, where it arrived on June 17, 1885. The pieces _____ (put) together and the opening ceremony took place in October the following year. Every year, the Statue of Liberty _____ (visit) by many people from all over the world.

Cloze activity B

Miss Lin's Day at Work

Read the passage and fill in the blanks with the correct form of the verb in the brackets.

Miss Lin is a teacher in a kindergarten. Yesterday, Miss Lin worked from 7:00 am to 6:30 pm. It was a very busy day.

At 7:30 am, a new parent came in to register her child. When the parent left, the child started screaming. Usually, the parent stays for a few minutes if it _____ *(be)* the child's first day. However, the parent had to leave, so Miss Lin had to calm down the child by herself.

At 8:00 am, another parent dropped off her boy. Miss Lin noticed that the boy was coughing and sneezing. She asked the mother if the child _____ (*be*) sick. If a child _____ (*be*) sick, he should _____ *(stay)* home. The mother insisted that her son wasn't sick. Miss Lin let the boy stay, but she had to watch him carefully the rest of the day.

At 3: 30 pm, Miss Lin began to set up for the afternoon snack. After the snack, Miss Lin saw red spots on the new child's face and his face was puffy. If a child has allergies, the parent must _____ *(tell)* the staff on the first day. However, this new child's parent didn't.

Later in the playground, one of the children fell off the climber and cut his lip, so Miss Lin took the child inside and gave him a cold pack for his lip. When the child's mother picked the child up, she was very upset and told Miss Lin, "If my child _____ *(get)* hurt, you should call me immediately."

Appendix G

Instructions for Stimulated Recall

Before the stimulated recall starts, the interviewee will be informed that

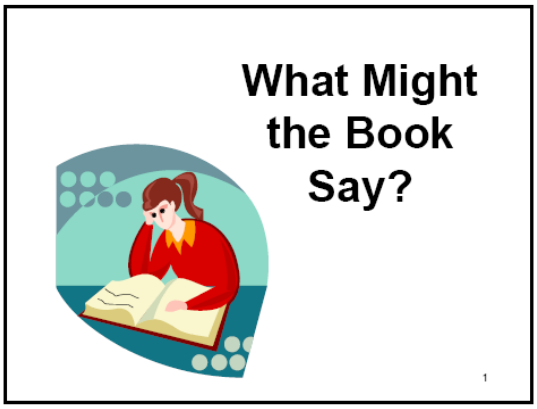
1. You can refuse to answer any question if you are not comfortable responding it.
2. The stimulated recalls will be audio-recorded and transcribed later for the data analysis. You can request to stop the recorder anytime you feel uncomfortable being recorded.
3. You can be assured that the stimulated recall data will be kept confidential and used only for the research purpose of this study.

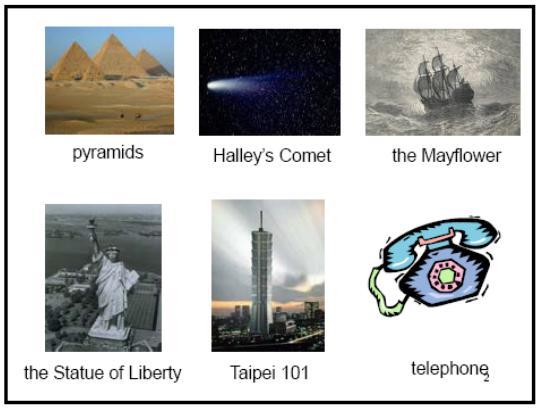
After the participant finished one cloze activity, he/she is to be asked:

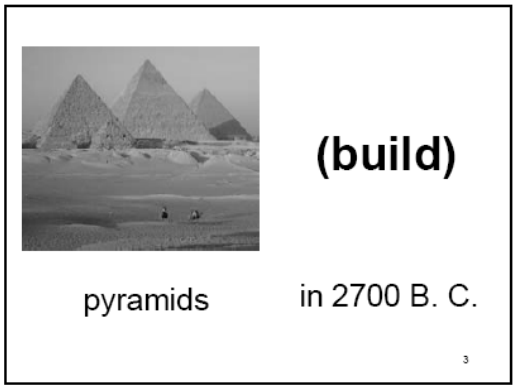
- Now you have finished the exercise, would you please tell me what you were thinking while you were doing the exercise?
- What you were thinking when you were doing item X?
- You said XXX just now. Could you tell me more about it?
- Did you mean X?
- Were you thinking X?

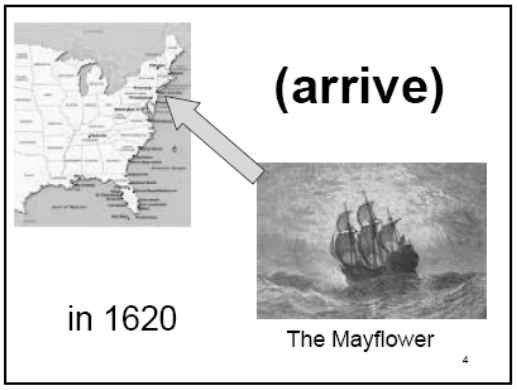
Appendix H

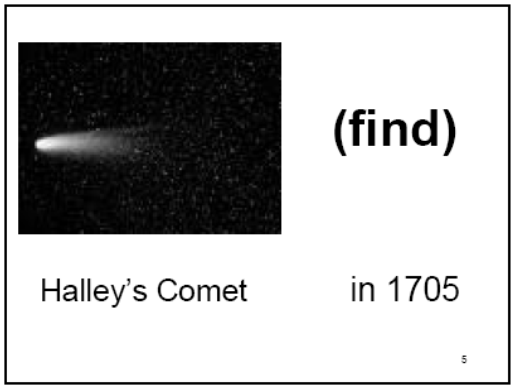
Oral Production Task (1) Targeting the Passive Construction

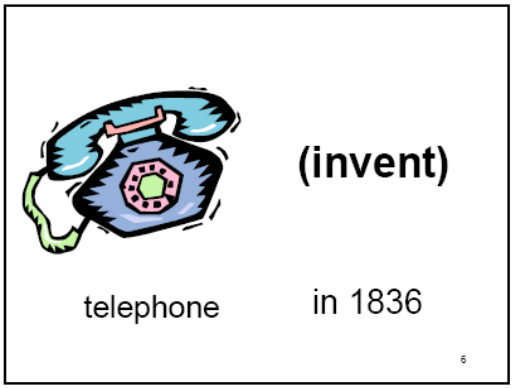


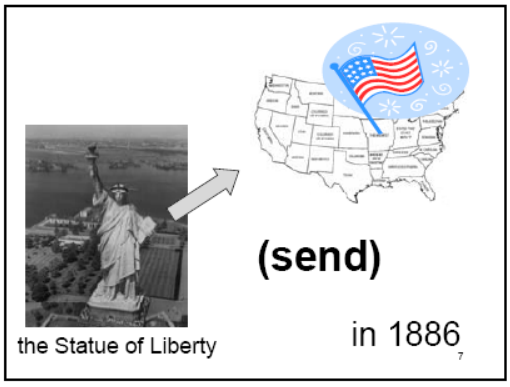








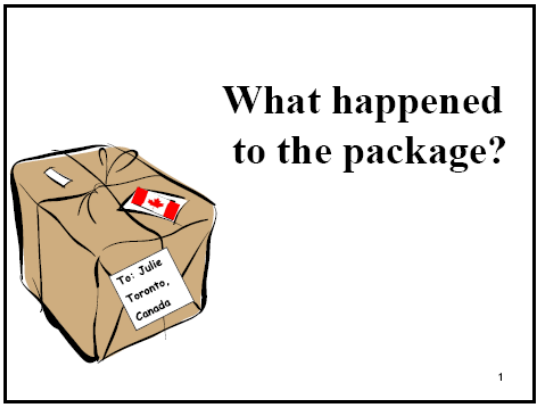


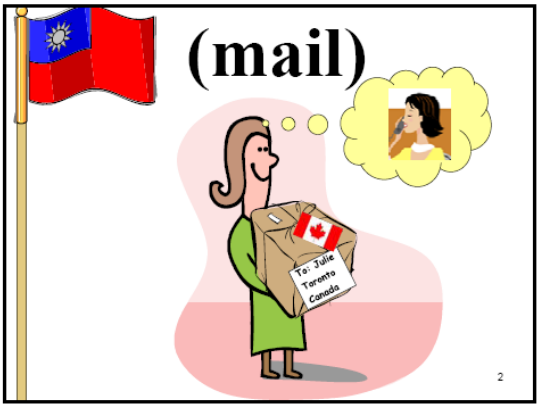

the Statue of Liberty
(send)
in 1886
7

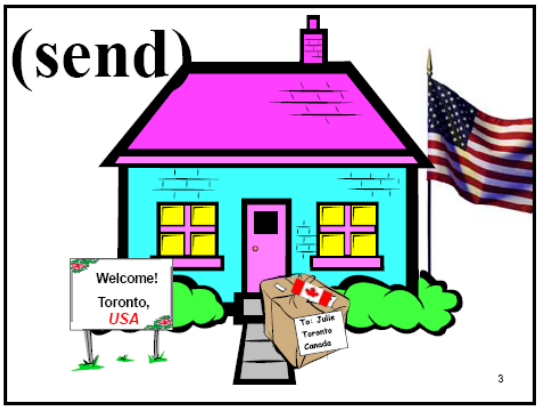

Taipei 101
(complete)
in 2003
8

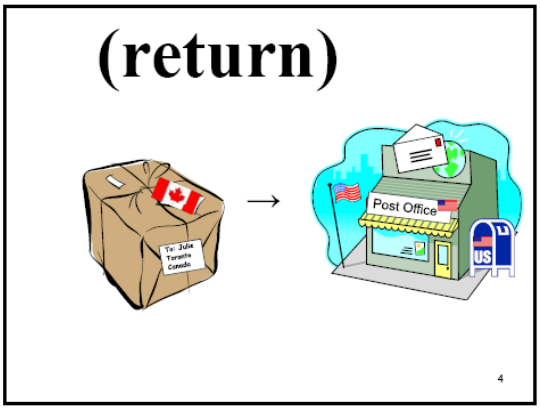
Appendix I

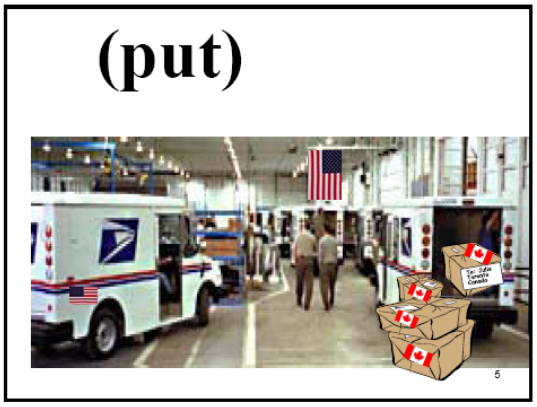
Oral Production Task (2) Targeting the Passive Construction

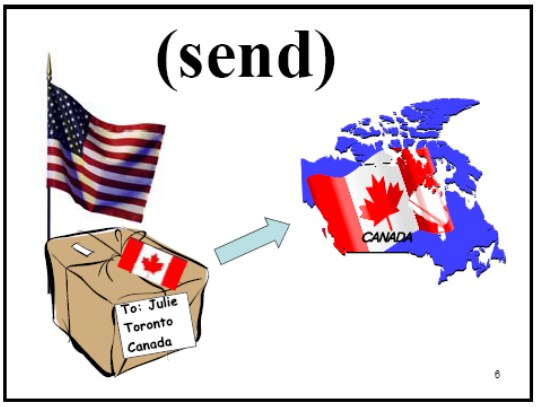




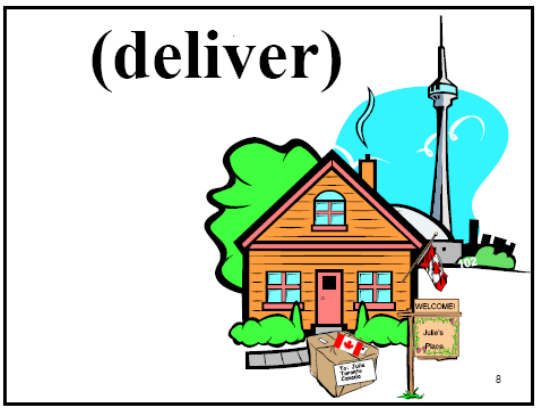


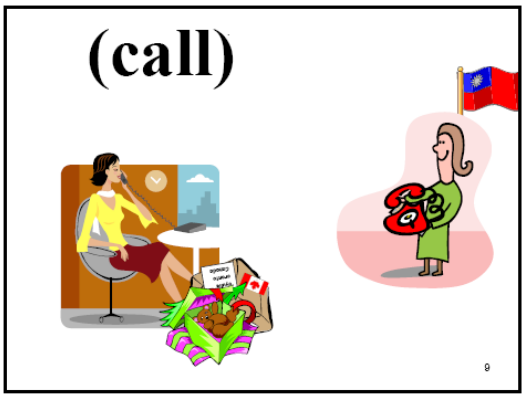






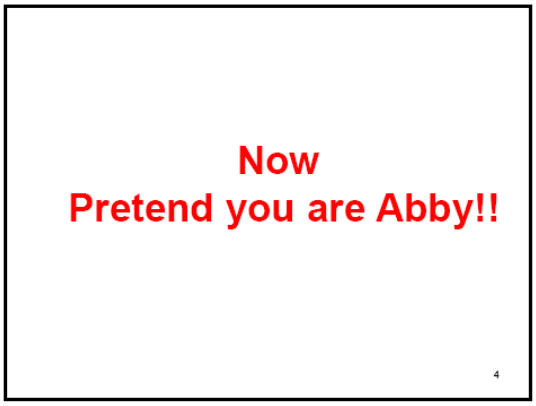

(arrive)
TORONTO
7

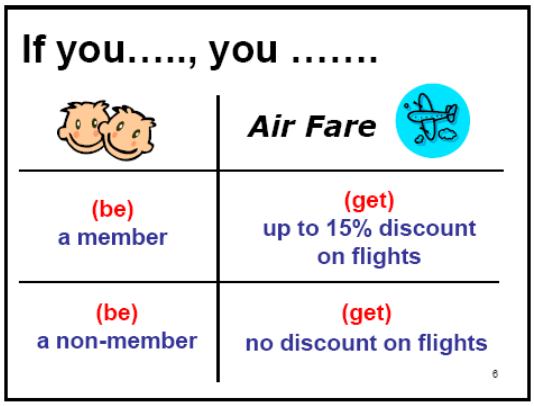

(deliver)
WELCOME!
8


(call)
9

Appendix J

Oral Production Task (3) Targeting the Real Conditional





If you….., you …….

	***Air Fare*** 
(be) a life-time member	(receive) 15% discount
(be) a one-year member	(receive) 5% discount

7

If you….., you …….

Hotel 	***Room Rate $$$***
(stay) at the River Hotel	(pay) only 80 dollars per night
(stay) at the Ocean Hotel	(pay) only 100 dollars per night

8

If you….., you …….

	***A thank-you gift***
(decide) today	(get) a free MP3 player

9

Appendix K

Written Metalinguistic Task

Instructions:

This error correction test consists of 40 items. Each of the items on the next pages contains ONLY ONE mistake. The mistake is not in the time frame of the sentence. Identify the mistake, correct it and explain your correction. You can write your explanation in either Chinese or English.

For example,

1. Everybody know that teenagers like to play computer games.

   The ungrammatical part is *know*

   The correct form should be *knows*

   It is ungrammatical because *"Everybody" is a singular noun subject, which needs to be followed with a singular verb form"*

2. The book give to John yesterday.

   The ungrammatical part is *give*

   The correct form should be *was given*

   It is ungrammatical because *"The subject noun phrase "The book" is a receiver of an action (give), which requires a passive verb form following it"*

Name: _____ **Date: _____**

1. They drive to Taipei twice last week.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

2. Many messages were receiving yesterday.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

3. If he taking the train, he will get to work by 8:00.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

4. My father work in Taipei.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

5. The test will be give orally.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

6. If my parents will come to visit me, I will take them to Taipei.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

7. Where you live?

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

8. Rules are making by the school.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

9. If my son goes to bed late tonight, he is tired tomorrow.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

10. Today, many products made of plastic.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

11. The twins crying if their mother leaves the room.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

12. He is tall than Mary.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

13. If it will snow tomorrow, they will cancel the meeting.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

14. The houses on this street built last year.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

15. My family will go to the zoo if the weather will be nice tomorrow.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

16. Millions of dollars have borrowed by developed countries.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

17. He like studying English.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

18. Many cars stolen last year.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

19. If we read to our children, they learning to enjoy books.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

20. The contract must signed by the boss.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

21. If it's too hot, the children should stayed inside.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

22. Many people have saved since insulin was discovered.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

23. Your book is cheap than mine.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

24. Laws are make by the government.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

25. Driving carefully if there is ice on the road.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

26. His heart has been breaking three times this year.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

27. Where you want to go?

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

28. If children eat breakfast, they having more energy.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

29. Vitamins were discovering in the early 1900's.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

30. Your proposal may be reject if you do not submit it by the deadline.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

31. My mom walk to work every day.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

32. My sister will build a snowman if there will be enough snow tomorrow.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

33. You must taken a taxi, if you miss the bus.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

34. If I using the computer too much, my eyes will hurt.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

35. What is they doing?

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

36. Children might start a fire if they playing with matches.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

37. The photocopier has been fix three times this year.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

38. If you walk away from the group, you might got lost.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

39. In France, poor health seen as a serious problem.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

40. The yellow dress is most expensive than the blue dress.

The ungrammatical part is _____

The correct form is _____

It is ungrammatical because _____

_____

_____

Appendix L

Table L1. *Spearman Correlations Among the 20 Grammar Structures*

	Present Perfect	Simple Past	Negation	Modal Verbs	Countable Noun	Passives	Articles	Unreal Con-ditionals	Embedded Questions	Third Person	Clauses	Present Progressive	Prepo-sitions	Adj Compara-tives
Present perfect	1.00													
Simple past	.54(**)	1.00												
Negation	.33(**)	.59(**)	1.00											
Modal verbs	.41(**)	.52(**)	.48(**)	1.00										
Countable noun	.35(**)	.37(**)	.31(**)	.36(**)	1.00									
Passives	.55(**)	.48(**)	.36(**)	.45(**)	.49(**)	1.00								
Articles	.40(**)	.42(**)	.37(**)	.38(**)	.52(**)	.46(**)	1.00							
Unreal conditionals	.43(**)	.24(**)	.18	.27(**)	.36(**)	.53(**)	.31(**)	1.00						
Embedded questions	.39(**)	.30(**)	.31(**)	.40(**)	.28(**)	.52(**)	.33(**)	.55(**)	1.00					
Third person –s	.28(**)	.47(**)	.56(**)	.42(**)	.41(**)	.41(**)	.40(**)	.22(**)	.27(**)	1.00				
Clauses	.41(**)	.38(**)	.33(**)	.34(**)	.31(**)	.51(**)	.33(**)	.49(**)	.61(**)	.30(**)	1.00			
Present progressive	.36(**)	.56(**)	.48(**)	.46(**)	.37(**)	.47(**)	.47(**)	.29(**)	.35(**)	.44(**)	.38(**)	1.00		
Prepositions	.34(**)	.32(**)	.28(**)	.33(**)	.52(**)	.42(**)	.43(**)	.43(**)	.36(**)	.33(**)	.41(**)	.39(**)	1.00	
Adjective comparison	.39(**)	.43(**)	.37(**)	.32(**)	.50(**)	.45(**)	.40(**)	.38(**)	.36(**)	.36(**)	.45(**)	.42(**)	.53(**)	1.00
Infinitives	.35(**)	.46(**)	.31(**)	.42(**)	.35(**)	.51(**)	.39(**)	.48(**)	.46(**)	.37(**)	.54(**)	.47(**)	.49(**)	.48(**)
Real conditionals	.41(**)	.27(**)	.21(**)	.24(**)	.28(**)	.42(**)	.28(**)	.71(**)	.48(**)	.23(**)	.47(**)	.30(**)	.43(**)	.41(**)
Wh-questions	.30(**)	.47(**)	.36(**)	.39(**)	.34(**)	.46(**)	.45(**)	.34(**)	.37(**)	.42(**)	.49(**)	.47(**)	.43(**)	.49(**)
Participial construction	.39(**)	.37(**)	.26(**)	.34(**)	.32(**)	.51(**)	.30(**)	.56(**)	.56(**)	.25(**)	.57(**)	.30(**)	.41(**)	.48(**)
Question tags	.33(**)	.45(**)	.37(**)	.36(**)	.27(**)	.40(**)	.27(**)	.42(**)	.47(**)	.36(**)	.46(**)	.42(**)	.32(**)	.41(**)
Past progressive	.42(**)	.48(**)	.33(**)	.35(**)	.31(**)	.44(**)	.39(**)	.31(**)	.40(**)	.40(**)	.42(**)	.52(**)	.36(**)	.38(**)

Note: ** Correlation is significant at the 0.0025 level (2-tailed).

	Infinitives	Real Condi-tionals	Wh-questions	Participial Construction	Question Tags	Past Progressive
Present perfect						
Simple past						
Negation						
Modal verbs						
Countable noun						
Passives						
Articles						
Unreal conditionals						
Embedded questions						
Third person –s						
Clauses						
Present progressive						
Prepositions						
Adjective comparison						
Infinitives	1.00					
Real conditionals	.57(**)	1.00				
Wh-questions	.50(**)	.39(**)	1.00			
Participial construction	.57(**)	.56(**)	.40(**)	1.00		
Question tags	.54(**)	.47(**)	.40(**)	.51(**)	1.00	
Past progressive	.50(**)	.40(**)	.49(**)	.41(**)	.54(**)	1.00

Appendix M

Transcription Conventions

[ ] Transcriber's commentary

< > Pauses of more than 5 seconds (e.g., <7>)

? Indicates rising intonation

Bold Emphasis

… Pauses of less than 5 seconds

– Incomplete utterance