Corrective feedback and Chinese learners' acquisition of English past tense

By

Yingli Yang

Department of Integrated Studies in Education

McGill University, Montreal

July, 2008

A Thesis Submitted to McGill University

in Partial Fulfillment of the Requirements of the

Degree of Doctor of Philosophy

Abstract

This quasi-experimental study investigated how recasts and prompts benefited Chinese learners' acquisition of regular and irregular English past tense. The research hypotheses were: 1) Learners who receive corrective feedback while performing the communicative activities will outperform those in the control group who are carrying out the communicative activity only; 2) Prompts will have more beneficial effects than recasts in learning regular past tense forms; 3) Recasts will better assist learners in learning irregular past tense forms than prompts.

Seventy-two adult EFL learners in 3 intact classes at the university level participated in the study. The 3 classes were randomly assigned to one prompt group (n = 22), one recast group (n = 25), and one control group (n = 25). The instructional treatment included four communicative tasks which highlighted the target forms and elicited oral production of the target forms from the students. In the two feedback groups, teachers consistently provided one type of feedback (i.e., either recast or prompt) in response to learners' errors during the communicative activities, whereas in the control group the teacher provided feedback only on content of the communicative tasks. Participants' acquisition of the past tense forms was assessed through an oral and a written test prior to, immediately after, and two weeks after the treatment.

Results of mixed-model repeated measures ANOVA showed that all three groups significantly improved their accuracy scores over time in both oral and written production. Results of the written production measure indicated that the prompt group outperformed

the control group and this improvement resulted from significantly more accurate use of irregular past-tense forms.

These findings indicate the beneficial role of form-focused communicative activities in L2 learning in form-oriented Chinese EFL classrooms and confirm findings in other previous studies, pertaining to the superior effectiveness of prompts over recasts in second language acquisition.

Résumé

Cette étude quasi-expérimentale a examiné les effets de deux techniques de rétroaction, reformulations et incitations, sur l'acquisition des verbes réguliers et irréguliers au passé (en anglais) par des apprenants chinois. Les hypothèses de recherche étaient les suivantes : 1) les apprenants qui reçoivent une rétroaction corrective dans le cours d'activités de communication seront plus performants que les membres du groupe témoin qui réalisent seulement les activités de communication ; 2) les incitations auront davantage d'effets bénéfiques que les reformulations dans l'apprentissage de verbes réguliers au passé ; 3) les reformulations seront plus avantageuses pour l'apprentissage de verbes irréguliers au passé.

Soixante-douze apprenants d'anglais langue étrangère dans trois classes complètes de niveau universitaire ont participé à l'étude. Les apprenants des trois classes ont été répartis, au hasard, en trois groupes : un groupe « incitation » (n = 22), un groupe « reformulation » (n = 25) et un groupe témoin (n = 25). L'expérience pédagogique comportait quatre tâches de communication qui mettaient en présence les structures ciblées et suscitaient la production orale de ces structures par les apprenants. Dans les deux groupes « rétroaction », les professeurs donnaient invariablement un type de rétroaction (soit reformulation, soit incitation) en réponse aux erreurs des apprenants dans le cours des activités de communication tandis que, dans le groupe témoin, le professeur ne fournissait de la rétroaction que sur le contenu. L'acquisition par les participants des verbes au passé a été évaluée au moyen d'un test oral et d'un test écrit avant, immédiatement après et deux semaines après l'expérience.

Les résultats de tests ANOVA montrent que les trois groupes ont, avec le temps, amélioré de manière significative la précision de leurs productions des formes au passé, tant en production orale qu'écrite. Les résultats de la mesure de la production écrite ont indiqué que le groupe « incitation » a mieux performé que le groupe témoin et que son amélioration s'est traduite par un usage notablement plus précis des verbes irréguliers au passé.

Ces résultats montrent les avantages des activités de communication centrées sur la forme dans l'acquisition d'une langue seconde dans des cours d'anglais langue étrangère en Chine et confirment les résultats d'autres études concernant l'efficacité supérieure des incitations par rapport aux reformulations en acquisition d'une langue seconde.

Acknowledgement

This research benefited from the support and help from a number of people. I would like to take this opportunity to show my sincere gratitude to each one of them.

I would like to express my heart-felt gratitude to Professor Roy Lyster, my thesis supervisor, who guided me through the different stages of my Ph.D study with insightful comments, constructive feedback, and most important of all, continuous support and encouragement. His integrity as a scholar also inspired me to a great extent and will benefit my academic career well beyond my years at McGill.

I would also like to thank my committee members, Professor Mela Sarkar and Professor Elin Thordardottir for their consistent support and guidance. I benefited a great deal from their knowledge of language acquisition and their thought-provoking feedback and suggestions. I am fortunate to have had the opportunity to learn from them.

Special thanks go to Dr. Jose Correa for his statistical assistance and Dr. Zhidong Zhang for generously spending time discussing statistical issues pertaining to this Ph.D project with me.

I am also indebted to professors and graduate students at McGill, for their extensive knowledge and scholarly advice that were constructive in the development of my Ph.D dissertation. I benefited a great deal also from discussions and seminars at the McGill Language Acquisition Program. Special thanks go to Professor Fred Genesee and fellow students at the seminar.

This dissertation would not have been completed without the generous support and help from the teachers, students, and administrative staff at the university where I

conducted the present study. I am truly grateful for their interest in the study and their co-operation during the data collection period. It is a regret that I can not disclose their names for ethical concerns.

I would like to express my deep gratitude to Rui Liu for her assistance in the data-coding process, Susan Ballinger for her help in the examination of the measurement material, Nootan Kumar for editing the dissertation, and Chen Qu for her generous help and great friendship. Thanks also go to my friends and fellow students in the Faculty of Education who provided moral support and shared their knowledge with me at different stages of my Ph.D study.

This dissertation is dedicated to my friends and my family. I owe a deep debt to my parents whose unconditional love and support encouraged me to finish my study; my husband, who never hesitated to provide moral and technical support in helping me complete this dissertation; my parents-in-law who made great sacrifices to help me; my sister and my daughter, whose love and smile were the intrinsic motivation throughout this process.

This dissertation was funded by Fonds Québecois pour la Recherche sur la Société et la Culture (No. de Dossier:120263) and McGill Internal SSHRC fellowship for dissertation research.

List of Tables

Tables in Text

Tables in Appendix D

Table of Contents

CHAPTER 1: INTRODUCTION

The empirical research described in this thesis investigates the effect of different types of feedback on the acquisition of English past tense forms and grew out of my interest as a teacher and as a researcher in the pedagogical choices of communicative language teaching in the Chinese context. Specifically, the research was inspired by the tension between a prevailing favorable view of the communicative teaching method, the task-based approach in foreign language teaching stipulated by the Chinese State Education Development Commission and the ensuing mixed attitudes of both teachers and students regarding the implementation of this approach in the language classroom (Rao, 2002).

Some researchers hold an absolutist view towards communicative language teaching, claiming that the adoption of this approach will bring about positive effects (e.g., Liao, 2004); others cast doubt on the effectiveness and the practical value of this approach (Bax, 2003; Hu, 2005). Prior to arguing whether communicative language teaching is applicable in a Chinese context, the other related questions that need to be answered are: (a) Which types of communicative teaching method would best fit the Chinese context? (b) Does communicative language teaching imply a total abandonment of the teaching and practice of grammar? (c) What approach, then, would assist in the development of communicative competence without necessarily jeopardizing accuracy? All of these questions boil down to a balanced approach of integrating grammar instruction into communicative activities, an approach that is referred to in the second language literature as form-focused instruction (Lightbown, 1998; Spada, 1997).

1.1 Form-focused Instruction

Form-focused instruction refers to "any planned or incidental instructional activity that is intended to induce language learners to pay attention to linguistic form" (Ellis, 2001, pp. 1-2). Form-focused instruction differs from the traditional grammar-and-translation method in that the former relates these forms to their communicative functions; therefore, the forms noticed during communicative interaction may be more likely to be retrieved when confronted by similar communicative contexts. In contrast, the grammar points presented in the traditional decontextualized manner may be remembered in similar contexts such as a discrete point grammar test, but difficult to retrieve in a communicative context (Lyster, 2004a). In such a context, learners tend to treat language instruction as separate from language use (Lightbown, 1998).

The mutual exclusion of grammar and communicative contexts may explain why many learners in the foreign language context are able to achieve high scores in discrete point grammar tests, yet lack the ability to communicate fluently and accurately in communicative contexts (Hu, 2003). Hu (2003) demonstrates in his study that learning experience and regional differences may explain the difference in students' performance in high order language tasks and communicative competence. Rao (2007) summarizes common features of teaching and learning English in the Chinese context as: (a) concentration on intensive reading as a basis for language study; (b) use of memorization and repetition as fundamental acquisition techniques; (c) emphasis on grammar rules and linguistic details; (d) use of translation as both a teaching and learning strategy; and (e) the teacher's authority and the student's passive role. Within the highly form-oriented Chinese context, there is a need for the implementation of form-focused instruction

embedded in meaningful communicative activities in order to achieve a balance between the development of communicative competence and grammatical accuracy. One such form-focused instructional technique that has drawn researchers' attention in recent years is corrective feedback.

1.2 Feedback and Second Language Learning

The effect of corrective feedback in second language acquisition has provoked a substantial number of empirical studies as well as theoretical discussions in the past three decades. The nativists' language acquisition theory proposes that language learning is primarily input-driven, and feedback does not and should not play a significant role in language acquisition. This is based on their claim that the formation and restructuring of second language grammar is solely attributable to an innate human linguistic mechanism working in tandem with positive evidence. Accordingly, negative feedback has little impact on language learning, merely affecting performance but not leading to changes in underlying competence (Schwartz, 1993). A similar position in second language acquisition in the 1980s maintained that all learners needed to acquire a second language was exposure to comprehensible input and motivation to acquire the L2 (Krashen, 1985).

In contrast, other researchers have argued that corrective feedback (or negative evidence) facilitates second language acquisition by drawing learners' attention to errors in their interlanguage and assisting in their second language development (Chaudron, 1988; DeKeyser, 1998, 2001; Spada & Lightbown, 1993; Spada, 1997). Cognitive theories have also acknowledged the significant role that feedback plays in the "cognitive

comparison" between learners' interlanguage and the target language (Nelson, 1981, 1987; Tomasello & Herron, 1989) as well as in the process of automatizing partially acquired target features in the L2 (DeKeyser, 1998, 2001).

While it is generally contended in L2 literature that feedback has a positive effect on second language acquisition, recent studies in Canadian immersion and ESL contexts have shown that not all feedback types are equally effective (Ammar & Spada, 2006; Lyster & Ranta, 1997; Lyster, 2004b; Panova & Lyster, 2002), as measured by either "uptake rate" or "gains in test scores." The results of these studies contrast with the findings from a number of studies conducted in laboratory settings (Long, Inagaki, & Ortega, 1998; Mackey & Philp, 1998; Ishida, 2004).

As pointed out by Ammar (2003), some of the factors that have led to the discrepancies in previous research are differences in definitions and operationalizations of different feedback types. Therefore, before addressing the research questions, it is necessary to examine the definition and terminology of the important constructs in the present study.

1.3 Definition and Terminology

Drawing on Lyster's (2004b) classification, feedback moves were differentiated as one of two types (i.e., recasts or prompts) in the present study. The fundamental difference between recasts and prompts is that the former provides learners with the correct form, whereas the latter offers learners various cues to self-repair while withholding the target form. By providing recasts, the teacher supplies the target form in

a natural flow of conversation and implicitly reformulates part or all of the student's utterance, as shown in Example 1.1.

Prompts include a range of feedback types: (a) elicitation, in which the teacher directly elicits a reformulation from the student by asking questions such as "How do we say that in English?" or by pausing to allow the student to complete the teacher's utterance, or by asking the student to reformulate his or her utterance; (b) metalinguistic clues, in which the teacher provides comments or questions related to the well-formedness of the student's utterance such as "We don't say it like that in English"; (c) clarification request, in which the teacher uses phrases such as "What?" or "I beg your pardon?" following learner errors in order to indicate to students that their utterance is ill-formed in some way and that a reformulation is required; and (d) repetition, in which the teacher repeats the student's ill-formed utterances, adjusting intonation to highlight the error. According to Lyster and Mori (2006), prompts are pragmatically different from explicit correction and recasts in that "By prompting, a teacher provides cues for learners to draw on their own resources to self-repair, whereas by providing explicit correction or recasting, a teacher both initiates and completes a repair within a single move" (p. 272). Examples of the two types of feedback are as follows[1]:

Example 1.1

Recasts

S: Once upon a time, there live a beautiful girl named Cinderella.

T: Once upon a time, there **lived** a beautiful girl named Cinderella.

Example 1.2

[1] Unless otherwise indicated, the examples in the present study come from classroom transcripts of the present study.

Prompts (Repetition)

S: Mrs. Jones travel a lot last year.

T: Mrs. Jones **travel** a lot last year?

S: Mrs. Jones travelled a lot last year.

Descriptive studies in communicative and immersion contexts (Lyster & Ranta, 1997; Panova & Lyster, 2002) have shown that prompts were more likely to elicit student responses than recasts, and thus more conducive to noticing. However, other studies in EFL settings (e.g., Korean EFL, Sheen, 2004), including my pilot study in the Chinese EFL context (Yang, 2006), have claimed that language learners were primed to notice the corrective purpose of recasts as well as other types of feedback because of their form-focused language learning experience. Experimental and quasi-experimental studies that have compared the relative efficacy of prompts and recasts have also yielded mixed results. While some studies have demonstrated the superiority of prompts over recasts (Ammar & Spada, 2006; Ellis, Loewen, & Erlam, 2006; Ellis, 2007; Lyster, 2004b), others did not show such an advantage for prompts (Loewen & Nabei, 2007).

Interestingly, most of the previous studies that have professed prompts as superior over recasts targeted rule-based grammatical features. Until now, the exact nature and efficacy of different types of feedback (i.e., recasts and prompts) on various types of grammatical features as well as the cognitive mechanism they invoke still remain to be examined (Ellis, 2007). Research in this area is especially scarce in EFL contexts, where L2 pedagogy and even learning processes may differ to a large extent from that of immersion or content-based contexts. In addition, past tense has been shown

to be a form difficult to acquire for Chinese learners even at the advanced level partly because in Mandarin Chinese, past tense is not morphologically marked as in English (Cai, 2007). The present study investigates how different types of feedback may benefit Chinese EFL learners' acquisition of regular and irregular English past tense, in an attempt to answer the question of whether recasts and prompts have distinct functions in the acquisition of rule-based versus exemplar-based grammatical forms.

1.4 Research Questions

This study is designed to compare the efficacy of two feedback types operationalized as "recasts" and "prompts" on the acquisition of two types of target structures: regular (-ed) and irregular past tense forms in English. It is hypothesized that the two feedback types may have differential effects on the acquisition of rule-based versus exemplar-based structures. Specifically, recasts will favor the development of exemplar-based irregular past tense forms, while prompts will better facilitate the acquisition of rule-based regular past tense forms in English.

The study employed a quasi-experimental design, in which four intact EFL classes composed of university-level adult learners participated. During the treatment sessions, the feedback groups received feedback while performing communicative activities, whereas the control group, while performing the communicative activities, did not receive feedback on the target grammatical structure. The effects of corrective feedback on learning the target forms were assessed by means of oral and written tests which have been claimed to tap into both implicit and explicit second language

knowledge (Ellis, 2005).

The main objectives of the study are two-fold: a) to investigate the overall effect of communicative activities combined with feedback on EFL learners' acquisition of English regular and irregular past tense, and b) to compare the relative effect of recasts and prompts on Chinese EFL learners' acquisition of regular and irregular past tense in English.

Specifically, the questions addressed in the present study are as follows:

Research Question 1

Q1: Do the groups that perform communicative activities while receiving feedback show an overall superiority in learning regular and irregular past tense over the control group not receiving feedback?

Research Question 2

Q2: To what extent do prompts and recasts differ in facilitating EFL learners' acquisition of regular English past tense?

Research Question 3

Q3: To what extent do recasts and prompts differ in assisting EFL learners' acquisition of irregular past tense?

1.5 Thesis Outline

There are six chapters in this dissertation. The present chapter introduces the purpose and context of the study, as well as the research questions and research procedures.

Chapter 2 presents the theoretical framework related to feedback studies including second language theories and issues specifically pertaining to feedback, such as the classification and efficacy of feedback.

Chapter 3 reviews several empirical feedback studies, including descriptive studies that document patterns of corrective feedback and learner uptake and repair, as well as experimental and quasi-experimental studies that examine specific types of feedback and their efficacy in second language learning. Drawing on the results of previous studies and the theoretical framework presented in Chapter 2, research questions and hypotheses of the present study are outlined at the end of Chapter 3.

Chapter 4 describes the methodology adopted in the present study. Specifically, the research context, participants, procedures, as well as the treatment and testing material are presented in detail. At the end of Chapter 4, the coding procedure and inter-rater reliability issue are also addressed.

Chapter 5 presents the data analysis and results of the study. Results of analysis of classroom transcripts of the treatment sessions, questionnaire data, as well as quantitative analysis of learners' test performance in the three tests are presented and analyzed.

Chapter 6 discusses the results and analysis of the present study, interpreting the findings according to theoretical and pedagogical issues pertaining to feedback efficacy in relation to the grammatical structures. This chapter also points out the limitations of the study and suggests future directions of feedback studies. The chapter concludes with the summary and conclusion of the present study as a whole.

CHAPTER 2：THEORETICAL FRAMEWORK ON THE ROLE OF FEEDBACK IN SLA

The debate over the role of feedback in second language learning has been well documented in the second language acquisition literature as well as in various other disciplines, including psychology (e.g. de Bot, 1996), cognitive science (e.g. DeKeyser, 1998, 2001) and linguistics (e.g. Schwartz, 1993; White, 1987). Researchers' views vary to a great extent regarding the types of evidence provided by feedback, their classification, as well as their role in the development of specific domains of a second language.

Underlying these different views are diverse theoretical stances regarding the role of negative evidence and positive evidence (Gass, 2003), input and output (Krashen, 1982, 1985; Swain, 1985, 1995), and explicit and implicit learning (DeKeyser, 2001; Ellis, 1994; Schmidt, 1990,1995) in SLA, which reflect the researchers' academic background and the research context from which they come.

This chapter begins by reviewing linguistic concepts in the discussion of general mechanisms of second language acquisition relevant to the study of feedback. Following this review, issues specifically related to feedback classification and efficacy in the larger picture of second language learning processes are raised. Finally, the discussion of different theories and hypotheses that are conducive to the role of corrective feedback in second language learning is presented.

2.1 The Role of Input and Output in SLA

In second language acquisition research, there has been an on-going debate over the role of positive evidence and negative evidence as well as the effect of input and output on language learning.

2.1.1 Definition of Types of Evidence

There is a general consensus in both L1 and L2 literature that language learners are exposed to two types of input, namely, positive evidence and negative evidence. According to Gass (2003), positive evidence refers to the input that consists of a set of well-formed sentences, speech samples that are available from the spoken language and/or from the written language. These speech samples are referred to as the most direct means that learners have available from which they can form linguistic hypotheses. Positive evidence has been noted to be the most obviously necessary requirement for both L1 and L2 acquisition (Gass, 2003).

Negative evidence refers to the type of input that is provided to learners concerning the incorrectness of an utterance. This might be in the form of explicit or implicit information. Based on Long and Robinson's (1998) taxonomy, negative evidence can be composed of two types: pre-emptive (occurring before an actual error — as in a classroom context), i.e. pre-planned, or reactive. If reactive, it can be explicit or implicit. Explicit negative evidence is an overt correction. Implicit negative evidence can take the form of either a communication breakdown or a recast. While Long and Robinson's taxonomy of negative evidence clearly defines recasts as an implicit type of negative

evidence, other scholars argue that this may not necessarily be true in all instructional contexts (Ellis & Sheen, 2006; Lyster & Mori, 2006).

In the present study, feedback is an umbrella term which includes a variety of techniques ranging from the most explicit to the most implicit end of the continuum. The operational definition of feedback in the present study is a teacher's or interlocutor's reaction which points out the incorrectness of the learner's utterance occurring in both naturalistic environments (i.e., in conversational interaction) and instructed environments (i.e., in the classroom).

The distinction among types of evidence has theoretical importance for language acquisition. It is not difficult to understand that positive evidence is the most obviously necessary requirement for learning since one must be exposed to a set of grammatical sentences in order for learning to take place. However, the role of negative evidence in second language acquisition has continued to be a controversial issue over the past three decades. The evolution of the evidence as well as the input versus output debate is discussed in further detail below.

2.1.2 Krashen's Input Hypothesis

The Input Hypothesis put forward by Krashen (1982, 1985) dismisses all traditional classroom interventions (e.g. grammar teaching and error correction) and the role of negative evidence in language learning. Krashen (1982) has claimed that all second language learners need to acquire language is exposure to sufficiently rich comprehensible input; that knowledge of consciously learned language is distinct from

unconsciously acquired language in representation; that only the latter type of knowledge can be deployed in spontaneous language use and, furthermore, that there can be no interaction between "learning" and "acquisition." The non-interface position states that learned knowledge can never become acquired knowledge.

However, other linguistic theories challenged the nativists' view and Krashen's Input Hypothesis which both emphasize the role of input and positive evidence in language learning. For example, White (1987) points out the importance of feedback, particularly as a source of negative evidence to indicate the inadequacy of learners' interlanguage system, and suggests that positive evidence alone is insufficient in second language acquisition. She further proposes that it is negative evidence triggered by incomprehensible input that becomes the impetus for learners to recognize their inadequate rule system, especially when they are required to go from a broader grammar (superset) to a narrower grammar (subset).

2.1.3 Swain's Output Hypothesis

Swain's (1985, 1995) Output Hypothesis challenges the traditional assumption that input is the only necessary requirement for second language acquisition, and that output is only the end product of second language acquisition that does not have any significant function in language acquisition processes (e.g., Krashen, 1985). From her research experience in French immersion contexts, Swain found that, despite years of exposure to sufficiently rich comprehensible input in communicative classrooms, students lacked grammatical accuracy in terms of morphology and syntax (Harley &

Swain, 1984; Lightbown & Spada, 1990, 1994). She therefore argued that comprehensible input alone is not enough for learners to produce grammatical and error-free utterances. Furthermore, one of the most important reasons for promoting output as a means to improve second language learning is that when learners experience communication difficulties, they need to be pushed into making their output more precise and appropriate.

In general, the importance of output in learning may be construed in terms of the cognitive processes triggered by output and learners' active engagement in these processes. As Gass, Mackey, and Pica (1998) pointed out, it is the necessity for learners to engage in syntactic processing to strive towards comprehensibility rather than comprehension of interlocutor input that may play a pivotal role in the acquisition process. It is also claimed that producing the target language may serve as "the trigger that forces the learner to pay attention to the means of expression needed in order to successfully convey his or her own intended meaning" (Swain, 1985, p. 249). Furthermore, Swain attributes considerable importance to corrective feedback, in effect associating the inadequacy of feedback as well as the inadequacy of pushed output with students' poor performance in grammatical accuracy in immersion classrooms.

Since the Output Hypothesis was first proposed, Swain has refined her hypothesis and specified the following four functions of output (Swain 1993, 1995, 1998). First, output has a fluency function which provides learners with opportunities for developing speedy access to their existing second language knowledge in the actual use of grammar in meaningful contexts. Second, output has a hypothesis-testing function. In the process of producing output, learners are able to form and test their hypotheses about

the comprehensibility and linguistic accuracy of their utterances in response to feedback obtained from their interlocutors. Third, output has a metalinguistic function. It is claimed that "as learners reflect upon their own target language use, their output serves a metalinguistic function, enabling them to control and internalize linguistic knowledge" (Swain 1995, p.126). In other words, output processes enable learners to reflect upon their use of the target language and consolidate their linguistic knowledge about the grammatical features of which they already have declarative knowledge. Reflection on language may enhance their awareness of forms, rules, and form-function mapping in a meaningful context. Finally, output serves as a noticing function. Namely, in producing the target language, "learners may notice a gap between what they want to say and what they can say, leading them to recognize what they do not know, or know only partially" (Swain, 1995, pp.125-126). The recognition of problems may then prompt the learners to selectively attend to the relevant information in the input, which will trigger their interlanguage development.

In summary, Swain's Output Hypothesis claims that output can, under certain conditions, promote language acquisition by allowing learners to invite feedback from interlocutors and teachers which informs them of the comprehensibility and well-formedness of their interlanguage utterances, also allowing them to move beyond semantic processing to syntactic processing. It is this syntactic processing that is believed to be an important element in the underlying second language acquisition mechanism.

2.1.4 Psycholinguistic Rationale for the Output Hypothesis

The widely held view on the role of output in second language acquisition has generated an increase in empirical investigations (Ellis & He, 1999; Kowal & Swain, 1994; Izumi & Bigelow, 2000, 2001) as well as in theoretical discussion (de Bot, 1996; Izumi, 2003).

Taking the information processing approach as the starting point, de Bot (1996) argues that output serves an important function in second language acquisition, specifically because it can generate highly specific input that the cognitive system needs in order to build a coherent set of knowledge. It is also claimed that from an information processing point of view, output plays a direct role in the process of turning declarative knowledge into procedural knowledge. Furthermore, de Bot (1996) argues that when the learner's output does not match the correct form, negative feedback will allow the learners to pay attention temporarily to language form instead of meaning, which would hamper the registration of the erroneous form in memory. In other words, on the one hand, output invites feedback that promotes noticing. On the other hand, feedback plays an indispensable role in pushing learners to produce grammatically more accurate output, which may consolidate already-learned knowledge of the rules, enhance form-meaning mapping, or trigger faster access to the already-learned structure to develop automaticity.

Interlocutors' feedback may also create the condition for cognitive comparison that enables learners to notice the gap between their interlanguage output and the target language input (Izumi, 2003). In a word, output and feedback are interrelated constructs that allow learners to selectively attend to useful information in the input and to modify

their ill-formed utterances, which are considered important processes in second language learning. The role of output may thus be seen as facilitating the process of input becoming intake.

2.2 Attention, Awareness, and the Noticing Hypothesis

The debate over positive evidence and negative evidence as well as input and output reflects one of the most controversial issues in applied linguistics, which concerns the role of conscious and unconscious processes in second language learning. On the one hand, there are many who believe that conscious understanding of the rule of the target language is necessary if learners aim to produce correct forms and use them in appropriate contexts. In this view, errors are the result of not knowing the rules of the target language, forgetting them, or not paying attention (Schmidt, 1990). Others firmly believe that language learning is an essentially unconscious process (e.g., Krashen, 1985). If second language learning were truly an unconscious process, then all second language learners should learn the language equally well in a rich second language environment.

2.2.1 Schmidt's Noticing Hypothesis

Drawing on data from the study by Schmidt and Frota (1986), Schmidt found that neither of the aforementioned two accounts could explain his own experience in learning Portuguese. Journal notes indicated that the forms he actually used were those he noticed people saying to him. Also, he found that a particular verb form that had been

taught did not guarantee that it would appear in his output. Presence and frequency of input did not account for what was actually learned. Furthermore, he found that only the linguistic forms that he noticed were incorporated into subsequent language output. Schmidt (1990) maintains that this study provides strong evidence for a close connection between noticing and emergence in production. This claim developed into one of the most influential theories in the second language acquisition field— the Noticing Hypothesis.

In Schmidt's Noticing Hypothesis, two important constructs in cognitive psychology were raised. Attention and awareness are thought of as accounting for the creation of new knowledge and/or the modification (restructuring) of existing knowledge. Attention and awareness are related, but not synonymous, constructs that subsume various concepts.

Schmidt (1995) points out that it is difficult to distinguish between attention and awareness. Attention is a limited-capacity system that is sometimes associated with the differentiation between controlled and automatic activities. Tomlin and Villa (1994) divide attention into three components: detection (the cognitive registration of stimuli), alertness (general readiness to deal with incoming stimuli), and orientation (the direction of attentional resources to specific stimuli). They argue that detection is necessary for learning because "detection is the process by which particular exemplars are registered in memory and therefore could be made accessible to whatever the key processes are for learning" (pp. 192-193). However, they believe that awareness may enhance alertness and orientation, yet detection does not necessarily imply awareness (Tomlin & Villa,

1994, p.198). It then follows that learning can take place without awareness but not without detection.

In contrast, Schmidt (1990, 1995) and Robinson (1995, 2003) argue for a critical role of awareness in learning. Robinson defines noticing as "detection with awareness and rehearsal in short memory" (1995, p.318) and distinguishes noticing from detection that is not accompanied by awareness. According to Robinson (1995), noticing the form of input is the result of attentional allocation leading to detection and rehearsal in short-term memory (STM), which is a necessary stage in SLA. Furthermore, what is noticed may be subsequently transferred to long-term memory (LTM). Noticing, therefore, plays an important role in illustrating the relationship between attention and memory.

Awareness is a subjective experience and commonly equated with consciousness (Schmidt, 1990). In his early work, Schmidt strongly opposes any unconscious learning, rejecting a dissociation of awareness and learning. It is posited that learners must consciously notice input in order for it to become intake. Schmidt differentiates two levels of awareness: awareness at the level of noticing (e.g. simply being aware of linguistic forms in the input) and awareness at the level of understanding (e.g. understanding the underlying rules of the linguistic form). It is awareness at the level of noticing that Schmidt claims is crucial for language learning, whereas awareness at the level of understanding is facilitative but not necessary for second language acquisition.

2.2.2 Conditions for "Noticing" and the Role of Instruction

By analyzing experimentally controlled studies, Schmidt (2001) claims that noticing is a necessary condition for storage of new forms; in other words, memory requires attention and awareness. He maintains that there are certain factors that determine what is noticed: (1) expectations, (2) frequency, (3) perceptual salience, (4) skill level, and (5) task demands. Schmidt proposes that, other things being equal, the more frequent a form is, the more likely it is to be noticed and then become integrated in the interlanguage system. Presumably this is because repeated presentation may enhance the opportunity for a form to have been noticed at some time or other. If attentional resources are variable and limited, then forms which are perceptually salient will have a greater chance of impinging on consciousness (Skehan, 1998).

In this respect, he argued, instruction may play an important role (Schmidt, 1990; Schmidt & Frota, 1986). Instruction can work in a more complex way by making the less obvious aspects of the input more salient, so that when the learner does the extraction and focusing, he or she may be prompted to notice certain features or parts of the target language in the input, which will in turn determine which part of the input can become intake subsequently. In a sense, learning is still input-driven (since the input is not being transformed) but in this way, instruction can allow the learner to choose what to prioritize in the input.

2.3 Feedback and its Functions in Second Language Acquisition

Feedback, as a reactive form of form-focused-instruction (Lightbown, 2001), has been claimed to be effective in promoting noticing, and thus conducive to second language learning (Mackey & Philp, 1998; Philp, 2003; Sheen, 2007; Trofimovich, Ammar, & Gatbonton, 2007). However, it has also undergone a substantial amount of discussion over the past three decades regarding its explicitness and/or implicitness, its effectiveness on second language acquisition, the type of evidence it provides, and more importantly, the learning mechanism it triggers (Ammar & Spada, 2006; Ellis & Sheen, 2006; Ellis, Loewen, & Erlam, 2006; Lyster, 2004b; Lyster & Mori, 2006; Sheen, 2007). The detailed discussion of the term feedback begins with a look at its classification.

2.3.1 Feedback Classification

Researchers in second language acquisition tend to differentiate feedback in terms of how explicit or implicit it is. In the case of implicit feedback, there is no clear indication that an error has been committed or where the error is, whereas in explicit feedback types, there is such an indication (Ellis et al., 2006). Implicit types of feedback often take the form of recasts, defined by Long (2007) as:

> a reformulation of all or part of a learners immediately preceding utterance in which one or more non-target like (lexical, grammatical, etc.) items are replaced by the corresponding target language form(s), and where, throughout the exchange, the focus of the interlocutors is on meaning not language as an object. (p. 2)

However, this dichotomous classification of feedback can be problematic. Research shows that depending on contexts (cf. Sheen, 2004) as well as characteristics of recasts

(such as linguistic targets, length, and number of changes), recasts can also be quite explicit, for example, by adding intonation and stress (Egi, 2007a).

Explicit types of feedback can also take a variety of forms according to different scholars. Ellis et al. (2006) claim that explicit feedback takes the form of either explicit correction, in which the response clearly indicates that the learner produced an erroneous utterance, or metalinguistic feedback, defined as "comments, information, or questions related to the well-formedness of the learner's utterance" (Lyster & Ranta, 1997, p.47). Others have operationalized explicit feedback in various ways. For example, Carroll (2001) and DeKeyser (1993) distinguished between explicit feedback that involved some specification of the nature of the error and explicit feedback that provides more detailed metalinguistic knowledge (also see Sheen, 2007).

Lyster (2002) questioned the reliability of comparing the effects of feedback in accordance with degrees of explicitness, based on his observation that it is self-repair which results from the illocutionary force of prompts rather than from their explicitness that contributes to second language development. He distinguishes recasts from prompts, which consist of clarification requests, repetitions, metalinguistic clues, and elicitation of the correct form. He draws such a distinction based on the idea that, while recasts provide learners with the target forms, prompts may engage learners in a deeper level of processing because they require the learner to retrieve the target form themselves. This view of categorizing different types of feedback is further illustrated in the following section.

2.3.2 Types of Evidence that Feedback Provides

While the explicitness of recasts and prompts is still the subject of heated debate (Ellis & Sheen, 2006), the theoretical argument on the type of evidence that different feedback techniques provide leads us to alternative approaches to differentiating feedback types.

As Nicholas, Lightbown, and Spada (2001) maintain, it is not difficult to understand that recasts afford learners with positive evidence, but whether they also provide negative evidence, as originally proposed by Long and Robinson (1998), is less clear. Other researchers (cf. Ellis & Sheen, 2006; Egi, 2007a) believe that whether recasts provide positive evidence, negative evidence or both, largely depends on learners' perceptions of them which in turn affect the effectiveness of recasts on language acquisition. If learners are consciously aware that they are being corrected, then they may perceive recasts as providing negative evidence, which may trigger the cognitive comparison between learners' interlanguage and the target language. If, on the other hand, in the meaningful interaction in which recasts occur, learners interpret recasts as a conversational reply confirming the content of the utterance rather than form, then recasts may simply serve as positive evidence (i.e., as examples of what is acceptable in the target language). The latter interpretation is highly likely in meaning-oriented classroom contexts (Lyster & Ranta, 1997; Panova & Lyster, 2002). Therefore, it is posited here that, depending on the context in which recasts are provided, they can provide positive evidence alone or positive and negative evidence concomitantly. In the latter case, it

follows that recasts take on a corrective characteristic whereas the former type of recasts provide positive exemplars only.

In the same vein, prompts, including repetition, clarification request, metalinguistic clues, and elicitation may also range from implicit to explicit depending on the discourse context. For example, Lyster (2002) argues that form-focused negotiations may appear relatively implicit if teachers in immersion classrooms feign incomprehension and provide simple prompts such as clarification requests. Therefore, he suggests that a more reliable way of differentiating feedback types would be in terms of whether or not they provide the correct reformulation because this would allow us to compare the effects of different retrieval processes.

As positive evidence in classroom input, recasts may help learners with establishing knowledge of new exemplars. Recasts elicit a relatively small number of modifications of students' ill-formed utterances and the uptake elicited may be only mechanical repetition of the correct form. Prompts, in contrast, clearly provide negative evidence to the learner since they signal that the learners' utterance is problematic as "a teacher provides cues for learners to draw on their own resources to self-repair" (Lyster & Mori, 2006, p. 272). The self-repair process is claimed to help learners re-analyze what they have already learned at some level and restructure their interlanguage (Lyster, 2002). According to de Bot (1996), learners benefit more from being pushed to "make the right connection on one's own" than from hearing the correct grammatical structures in the input (p. 549). Furthermore, prompts may help learners gain greater control over already-acquired forms and access them in faster ways. The discussion on the classification and relative efficacy of different feedback types reflects researchers' diverse perspectives on

how language learning takes place, and therefore would be better understood within a theoretical framework of second language learning mechanisms.

2.4 Theoretical Framework on the Relative Efficacy of Feedback

The debate concerning the efficacy of different feedback types has highlighted an underlying issue – that of the exact nature of the learning mechanisms involved in second language acquisition. Does explicit knowledge lead to better SLA than does implicit knowledge? Researchers have attempted to answer this question with both empirical studies as well as theoretical discussion. A detailed review of the various empirical studies conducted is provided in the following chapter. In this section the theoretical framework of SLA is discussed.

There are two other issues implied in the debate about the type of evidence that feedback provides, issues which involve the precise role and relative efficacy of evidence in the actual mechanisms and successful achievement of second language acquisition, respectively. With respect to feedback, the questions that remain to be answered are: a) do prompts contribute to the development of learners' implicit knowledge or do they enhance metalinguistic awareness (explicit knowledge) only; and b) if both types of feedback techniques can contribute to the development of implicit knowledge, is one type more effective than the other?

Empirical research aiming to compare the relative effectiveness of these two types of feedback have shown that the explicit type overall proves to be more effective in assisting the learning of certain grammatical structures (Carroll & Swain, 1993; Ellis et al., 2006; Lyster, 2004b). However, other studies in experimental and classroom settings

have also shown that recasts had positive effects on learning as well (Doughty & Varela, 1998; Long, Inagaki & Ortega, 1998; Mackey & Philp, 1998). Partly due to methodological discrepancies, studies on the effect of feedback yielded mixed results pertaining to the relative effect of the two types of feedback, yet they reflect the diverse theoretical frameworks upon which different scholars draw. These diverse theoretical frameworks are discussed in the following section in more detail.

2.4.1 The Interaction Hypothesis

Advocates of recasts mostly draw on the Interaction Hypothesis proposed by Long (1996), who claims that conversational moves such as recasts and clarification requests provide learners with a primary source of language input which enables them to negotiate meaning in a natural flow of conversation while facilitating language development. Long's Interaction hypothesis (1983, 1985, 1996) evolved from work by Hatch (1978) on the importance of conversation in the development of grammar, and from claims by Krashen (1985) that comprehensible input is a necessary condition for second language acquisition. In this view, during negotiation of meaning, when learners have communication breakdowns, they struggle to understand the interlocutors' utterance and to have the interlocutor understand their language, which leads to linguistic modification in a meaningful context. This allows them to pay attention to both meaning and form at the same time.

In Long's (1996) updated interaction hypothesis, he has pointed out the role of negotiated interaction which elicits negative feedback (including recasts) in second

language acquisition, "…negative feedback obtained in negotiation work or elsewhere may be facilitative of second language development" (p. 414). Recasts are highly valued as a feedback technique in this framework because they are assumed to provide both positive evidence and negative evidence by juxtaposing the correct and the incorrect utterances while keeping the meaning constant. In this way, they are thought to free up the learners' attentional resources which are limited and selective (Ammar & Spada, 2006).

On the other hand, many studies have indicated that those conversational moves should be classified as negotiation of meaning, since they may only provide learners and their interlocutors with some communicative strategies that facilitate comprehension instead of focusing on the accuracy of form (e.g., Lyster & Ranta, 1997; Pica, Young, & Doughty, 1987). Furthermore, Lyster (2004a) claims that there is little direct evidence that actually associates conversational moves used in negotiation of meaning with second language development. In addition, he questions the feasibility of negotiation of meaning in classroom settings where "…mutual comprehension of interlanguage forms reduces the need to negotiate for meaning" (p. 323).

An alternative theoretical framework for explaining second language learning is the information processing theory adapted from skill acquisition theory in cognitive psychology (de Bot, 1996; DeKeyser, 1998; Hulstijn, 1990; Lyster, 2004b). In the information processing framework, the role of feedback is prominent, since the procedualization of rule-based declarative representations results from practice and feedback (Lyster, 2004b), which enables learners to restructure their interlanguage. It is argued that without systematic feedback or other instructional interventions,

interlanguage representations can become fossilized and reach a developmental plateau. The information processing theory has been derived from Anderson's Adaptive Control of Thought (ACT) theory.

2.4.2 Anderson's ACT Theory

Anderson's adaptive control of thought (ACT) is a well-known general theory of skill acquisition which has stemmed from cognitive psychology (Anderson, 1983; 1985). ACT theory has been adopted by several second language researchers (de Bot, 1996; DeKeyser, 1998; Hulstijn, 1990; Lyster, 2004b) to explain processes of second language acquisition in particular. The information processing theory in second language acquisition assumes that language learning is a special case of skill acquisition which involves a gradual transition from declarative knowledge to procedural knowledge. Declarative knowledge (knowledge of "that") refers to consciously held, skill-relevant knowledge that is describable or factual knowledge. Procedural knowledge (knowledge of "how") is knowledge evident in a person's behavior but which the person is not consciously aware of and hence cannot describe in words (DeKeyser, 1998).

ACT theory maintains that, initially, the execution of a cognitive skill involves retrieving and using declarative knowledge which requires a great deal of learners' limited attention capacity. This is called *controlled information processing*. Then, through a process of *proceduralization* (the transition from declarative knowledge to procedural knowledge), which involves passing from a cognitive stage where learners use declarative knowledge extensively, and an associative phase where rules are applied repeatedly in a consistent manner, learners gradually arrive at an autonomous stage where

explicit knowledge of rules is no longer necessary. In this stage, learners may lose the declarative knowledge (although this is not necessarily the case, DeKeyser, 1998) but still can perform the task in a fast, coordinated fashion.

The information-processing model obviously has its strength in distinguishing two knowledge types and the developmental stages in language acquisition. However, since this model stems from cognitive psychology and applies to general skill acquisition rather than language *per se*, it inevitably has limitations in accounting for all aspects of the language acquisition process. For example, this model is established on the view that language is a rule-based representational system. It follows that the acquisition of language starts from explicit knowledge of the rules of the target language system, then, through gradual execution of the declarative knowledge, little attention is required in the actual deployment of the language. However, the assumption that language is a rule-based system has been challenged by a number of scholars (such as Skehan, 1998). There has been a debate regarding representational system in second language acquisition, both in the linguistic and psycholinguistic literature between proponents of a rule-based system (such as Reber, 1989) and advocates of what Skehan called "exemplar-based systems", i.e. accumulations of large numbers of formulaic items (such as the item-based theory proposed by Logan, 1988).

2.4.3 Alternative Theories of Language Representation

An alternative to Anderson's ACT theory of automaticity is Logan's (1988) instance theory or "item-based" theory. In this theory, the central idea is that

"Automaticity is memory retrieval: performance is automatic when it is based on single-step direct-access retrieval of past solutions from memory" (Logan, 1988, p. 493). The term "instance" here refers to a representation of co-occurring events or the events which are encoded in memory. According to this theory, which events are encoded in memory is determined by what learners pay attention to. There are three basic assumptions in this theory: a) Encoding into memory is an obligatory, unavoidable consequence of attention; b) Retrieval from memory is also an obligatory, unavoidable consequence of attention; c) Each encounter with a stimulus is encoded, stored, and retrieved separately. DeKeyser (2001) comments that Logan's instance theory is limited by the fact that only stimuli identical to the ones that were encountered before can be accessed in the retrieval process. This is not applicable in daily life, especially in the language learning process because the instances language learners are exposed to are somehow similar but certainly not all identical to previous examples.

Aside from Logan (1988), other researchers have also challenged the rule-based system hypothesis, arguing that such a view may over-emphasize the rule-governed basis of language (Bolinger, 1975). In reality, it is claimed that instances of language use are much more based on lexical elements. An additional perspective on the lexical aspect of language comes from empirical examination of large corpora of texts. For example, Sinclair (1991) argues that while grammar enables endless computational possibilities, in practice most such possibilities are ignored. Instead, particular combinations of lexical items occur again and again.

By reviewing previous literature on the representation of language, Skehan (1998) claims that both the psycholinguistic literature and linguistic literature coincide:

there is a possibility that the two systems co-exist, one being rule-based, the other being item-based. This can be illustrated by Palmeri's (1997) exemplar-based random walk model and Anderson, Fincham, and Douglas's (1997) experiment, which indicates a mixture of production rule use and item retrieval in the learning process. DeKeyser (2001) reviews previous research on the two different approaches and also comes to the conclusion that neither the rule-based system nor the item-based system could account for all the data in previous literature. Therefore, when it comes to a phenomenon as complicated as the process of second language acquisition, "an integration of the rule account and the similarity-based item-retrieval account will probably be necessary" (DeKeyser, 2001, p. 122). An example of an integrated rule and item-retrieval account is Skehan's Dual Mode System Hypothesis.

2.4.4 Skehan's Dual-Mode System Hypothesis

Drawing on previous theories in cognitive psychology and second language acquisition literature, Skehan (1998) proposed the Dual-Mode Hypothesis, claiming that the dual-code system (i.e., a rule-based system and an exemplar-based system) exists and functions at the same time:

> In any case, the linguistic and psycholinguistic arguments coincide here — the two systems co-exist, the rule-based analytic, on the one hand, and the formulaic, exemplar-based, on the other. In the former case, compact storage and powerful generative rules operate together to 'compute' well-formed sentences. In the latter, the central role is occupied by a very large, redundantly structured memory system, and (presumably) less powerful rules which operate on chunks much of the time, rather than on individual items. (p. 54)

He argues that the development of fluency depends, on the one hand, on the procedualization of the rule-based system with less control over the material, and on the other hand, on the memory-based chunks available for incorporation as wholes into language. This argument is based on the fact that the rule-based system is generative and flexible, but limited in terms of explaining fast on-line processing; while the exemplar (memory) system may be more rigid, but functions more quickly and effectively when communication demand is high. Skehan further argues that second language users seem to have the two systems at their disposal and move between the two systems quite freely. He hypothesizes that there are three stages in language acquisition in which the two systems work together. He argues that the initial stage of language acquisition is primarily lexical in nature (Nelson, 1981). "Contextualized coded exemplars are used to communicate meanings in a direct manner" (Skehan, 1998, p. 90). Then, at a later stage, processes of syntacticization come into play. Language learners begin to analyze the lexical-based input and organize the material into a rule-governed system. At a final stage, it is hypothesized that language which has been syntactized is then relexicalized. In other words, rules may be used to create exemplars that serve a communicative function.

The advantage of viewing language learning as a dual mode system is that it provides a rationale for a balanced structure in language teaching. In the meantime, the assumption of this hypothesis provides a strong interface position regarding the relationship between explicit and implicit knowledge, which underscores the role of instruction in the language learning process. In other words, how language learners acquire the target language largely depends on how language material is structured and how the learning process is manipulated. While maintaining that language learners have

to undergo these three stages, Skehan (1998) also acknowledges that the three stages are not necessarily sequential. In other words, on the one hand, learners need to be prepared to look for rules and identify patterns from the input, whereas on the other hand, the identification of rules is not the only purpose, and the analysis needs to be integrated to achieve fluent performance in real communicative contexts.

2.4.5 Pedagogical Implications of the Dual Mode System Hypothesis

Despite the fact that the Interaction Hypothesis and the Information Processing Theory have a number of different perspectives on second language learning mechanisms, they are not necessarily in direct contrast or mutually exclusive. Skehan's dual-mode hypothesis offers a convergence of the rule-based versus the exemplar-based view of language representation. DeKeyser (1998) also argues that the automatization viewpoint and the implicit learning viewpoint are not necessarily incompatible, as long as two points are considered. First is that the degree to which structures are most easily learned explicitly and which completely implicitly depends on the nature of the rule. Second, the term automatization can mean different things. Automatization is used either in the sense of fine-tuning or in a wider sense of restructuring, proceduralization, and fine-tuning. It is only in terms of proceduralization that the concept of automatization is incompatible with implicit learning, since proceduralization itself implies that rules start out as explicit.

In this view, it is unclear that implicit learning is totally impossible. In fact, DeKeyser argues that it depends on what type of language structure is at stake. For similarity patterns, learners most likely learn them by memorizing exemplars, whereas

for abstract rules they need more explicit pedagogical intervention. His classification of abstract rule versus similarity pattern depends on two factors: a) surface variation that tends to conceal the rule, and b) the distance between the co-occurring elements. These two factors make certain structures difficult to be perceived by the learners yet simple to state abstractly, and therefore require explicit focus on form. Take the rule for subject-verb agreement as an example. The plural noun and the plural marking on the verb can be separated by other elements such as adverbs and thus presents a difficulty for learners to induce the rule themselves. Yet the rule is stated abstractly rather easily: Whenever the subject is plural, mark the verb as plural. A similarity pattern, on the other hand, is of a different nature. For example, the prototypicality patterns in irregular past tense in English can be impossible to fully state abstractly, but if the patterns of co-occurrence of formal alternations and certain morphemes and allomorphs are predictable as a probabilistic, prototypical system, the provision of exemplars may enable learners to identify the pattern themselves. In this case, the implicit learning mechanism may work well.

In the same vein, Ellis (2006) discusses the criteria that determine the difficulty of linguistic structures, arguing that structures for which clear rules can be formulated (such as regular past tense) can be more easily learned as explicit knowledge than structures that necessarily involve item-learning (such as irregular past tense). He also relates this distinction to Hulstijn and De Graaf's (1994) distinction of "rule-learning" and "item-learning". Ellis's systematic analyses of seventeen grammatical structures measured by four tests convincingly support the view that distinguishes different knowledge types as well as acquisitional processes in language learning.

2.4.6 The Distinct Functions of Recasts and Prompts

If Skehan's Dual-Mode Hypothesis is tenable and if rule-based and item-based structures co-exist in the language system and involve various learning procedures, it follows, then, that Focus-on-Form activities that aim at distinct areas may serve different functions in second language learning. For example, DeKeyser (1995) found in his study that subjects learned simple abstract morphosyntactic rules in a miniature linguistics system significantly better under explicit learning conditions than under implicit conditions, but they learned similarity patterns better in the implicit inductive conditions than in the explicit-deductive conditions. Robinson and Ha (1993) also suggest that both structural complexity and developmental readiness may determine whether the learner relies on item-retrieval or rule-searching in their second language learning process.

In terms of the efficacy of different types of feedback on second language learning, Lyster and Mori (2006) maintain, "Prompts and recasts can be seen as complementary moves with different purposes for different learners in different discourse contexts" (p.273). They further hypothesize that recasts may provide exemplars of positive evidence and thus may be the right candidate for facilitating the encoding of new target representations when they occur in appropriate contexts, while prompts, because of their function as overt signals to elicit modified output without providing any positive evidence, may serve to enhance control over already acquired items by accelerating the transition of declarative to procedural knowledge (de Bot, 1996; Lyster, 2004b).

Although this hypothesis has theoretical foundations, it has yet to be empirically examined. The controversial issue of whether recasts can be an effective technique that offers negative evidence, and which type of feedback is more effective in assisting the learning of which types of grammatical features remain to be explored. Empirical data so far have either provided descriptive observations in the classroom where the effectiveness of recasts was measured by immediate student response only (Lyster & Ranta, 1997; Panova & Lyster, 2002; Sheen, 2004), or in experimental conditions where recasts appear to provide useful input in second language development (e.g. Long et al., 1998; Mackey & Philp, 1998).

The few classroom studies that actually compared recasts with prompts and that are comparable in methodology to some extent (Ammar & Spada, 2006; Ellis et al., 2006; Ellis, 2007; Lyster, 2004b) have shown an overall positive effect for prompts over recasts in facilitating second language development. Interestingly, most of these studies used language features that have been described by DeKeyser (1998) as rules that are abstract in nature — possessive determiners in English (*his/her*) in Ammar and Spada (2006), regular past tense *(-ed)* in Ellis et al. (2006), and regular past tense (*-ed*) and comparatives (*-er*) in Ellis (2007). It is perhaps because of the more abstract nature of the target language structures that the prompts were found in these studies to be more effective than recasts.

A target feature that is more difficult to categorize as unequivocally rule-driven, and therefore more difficult to associate with the appropriate feedback type, is grammatical gender in French. Contrary to many French grammarians and teachers who claim that gender attribution is arbitrary and unsystematic, Lyster (2004b) argues that

grammatical gender is largely rule-governed, but that the rules derive from similarity patterns based on word-internal properties (p. 408). The difficulty in categorizing grammatical gender as rule-based or exemplar-based highlights the difficulty in assigning the appropriate feedback type to the appropriate target language feature. Sheen (2007) notes in her study that, "Research has demonstrated that the noticing of the corrected feature in recasts depends largely on the linguistic feature that is being targeted" (p.319.). Ellis (2007) also argues that the effects of feedback would vary according to the structure being targeted and called for research that investigates how linguistic factors determine which different types of feedback will work for acquisition.

Based on a limited number of studies that compared recasts with prompts on a limited number of linguistic structures, it is speculated that most of the features that were tested in the studies favor prompts because they are rule-based; as a result, practice and prompts may assist learners to gain better control over or have faster access to these features. Contextual factors may also explain the overall positive effect of prompts, because most of these studies were carried out in either content-based or communicative ESL classes where meaning was the primary focus in the classroom. The question that needs empirical investigation is whether prompts would also work well when it comes to the learning of exemplar-based language features by EFL learners. In other words, can both types of feedback be effective in such a context, differing only in terms of the type of language structures they cater for? The answer to this question can contribute to larger theoretical issues regarding implicit and explicit learning, negative and positive evidence, and the cognitive model upon which feedback studies build.

Chapter Summary

This chapter reviewed previous literature relevant to the role of different types of evidence, input and output as well as cognitive theories in second language learning in search of a rationale for feedback as an appropriate focus-on-form technique in second language pedagogy. Specifically addressed was the question of what types of feedback cater to what types of learning.

It was argued that, in terms of linguistic input, both positive and negative evidence may contribute to second language acquisition in significant but perhaps different ways. Positive evidence provides learners with grammatically correct sentences that are processed by the learner to extract useful information (exemplar or abstract rules) for future language production. Negative evidence in second language literature has also been shown to be highly valuable, at least in certain aspects of second language development.

Different from the Input hypothesis which places the indispensable role of input in second language learning, Swain argues that output is complementary to input in allowing the learner to move from "semantic processing" prevalent in comprehension to more "syntactic processing" necessary for second language development. In the meantime, it was argued that in an interactive environment, feedback provides a tremendous resource for output to perform its four functions and to assist in the restructuring of interlanguage.

Cognitive theorists such as Schmidt believe that attention is a key concept in accounts of the development of L2 fluency (Schmidt, 1992). A number of researchers and

theorists have argued that there may be two types of learning (e.g., declarative and procedural, or explicit and implicit, or rule-based and instance-based) that differ in their reliance on awareness, but both depend on attention (Carr & Curran, 1994; Tomlin & Villa, 1994).

Drawing on previous theoretical frameworks in the study of feedback, it was argued that Skehan's (1998) dual mode system hypothesis complements previous models such as the Interaction Hypothesis and Anderson's information processing model. Skehan's model has a solid foundation in the psycholinguistic as well as the second language acquisition literature and is superior in explaining the mental state learners have at their disposal, as well as in illuminating the developmental stages that learners undergo.

Finally, the theoretical discussion led to a pedagogical issue concerning the role of feedback in second language learning, which has drawn the attention of researchers across various fields as well as teachers in real classrooms. Drawing on previous research, it has been argued that there is no clear answer pertaining to which type of feedback is superior to the other in learning grammatical structures. This is because the theoretical stances regarding learning mechanisms are still under debate, and more fine-tuned empirical research needs to be done before one can reach a conclusion.

However, based on the cognitive model proposed by Skehan (1998), it was hypothesized that recasts and prompts may serve different functions for learning different grammatical structures. Specifically, recasts favor the development of exemplar-based grammatical structures since they provide positive evidence that may lead to the registration of new exemplars or consolidate the partially acquired items, while prompts better facilitate the acquisition of rule-based structures since they require a deeper level

of processing through self-repair and therefore provide conditions for faster access to these structures. This hypothesis stems from cognitive theories of second language acquisition and needs to be tested empirically in both laboratory settings and classrooms for validity concerns. The next chapter reviews feedback studies conducted in a variety of settings, with a view to addressing methodological as well as contextual issues in relation to the efficacy of different types of feedback.

CHAPTER 3：EMPIRICAL STUDIES OF CORRECTIVE FEEDBACK

In the second language literature, a substantial amount of research has demonstrated that corrective feedback plays a facilitative role in second language development (Ammar & Spada, 2006; Lyster, 2004b; Mackey & Philp, 1998; Mackey, 2006). Numerous empirical studies have been conducted in different settings ranging from communicative and content-based immersion classrooms to form-oriented analytic foreign classrooms, with various data collection methods such as observational (Lyster & Ranta, 1997; Lyster & Mori, 2006; Panova & Lyster, 2002; Sheen, 2004), experimental or quasi-experimental (Ammar & Spada, 2006; Ishida, 2004; Long, Inagaki, & Ortega, 1998; Lyster, 2004b; Mackey & Philp, 1998; Muranoi, 2000), and stimulated recall methods (Egi, 2007a, 2007b; Kim & Han, 2007; Mackey, Gass, & McDonough, 2000; Mackey, Philp, Egi, Fuji, & Tatsumi, 2002; Mackey, 2006). These studies have revealed numerous facets of the issue, spawning a great deal of new perspectives on the function and effect of feedback in various contexts. This chapter will (a) review empirical studies of feedback with a view to comparing the effects of recasts and prompts; (b) discuss methodological and contextual issues in feedback studies, and (c) propose the rationale and hypotheses of the present study.

3.1 General Effect of Feedback in SLA

A plethora of empirical studies in second language research have well documented the role of corrective feedback in second language acquisition. My review of these empirical studies begins with studies that demonstrate the general effect of feedback, followed by observational studies describing patterns of feedback and uptake in different contexts, then proceeds with the discussion of the effect of recasts in relation to other feedback types, drawing on results from a series of experimental and quasi-experimental studies.

Many studies and meta-analysis on corrective feedback have reported the overall beneficial effects of corrective feedback on second language acquisition (Tomesello & Herron, 1988, 1989; Mackey, 2006; Russell and Spada, 2006).

In two classroom studies, Tomesello and Herron (1988, 1989) investigated the effects of feedback provided during teacher-led drills using the Garden Path technique[2]. They found that when teacher feedback enabled learners to engage in cognitive comparison between their own erroneous utterance and the target grammatical structure, the students learned better than in situations in which they simply received a series of correct exemplars of the new structure. Furthermore, their results could be applied to both L1 transfer errors and L2 overgeneralization errors.

Researchers have claimed that interactional feedback promotes L2 learning because it prompts learners' noticing of L2 forms. In order to empirically examine this issue, Mackey (2006) explored relationships between feedback, noticing, and subsequent L2 development. The target features were questions, plurals, and past tense forms.

[2] In the Garden Path condition, students' L1/L2 transfer errors are elicited intentionally and then corrected by their teachers.

Twenty-eight high-intermediate level adult ESL learners at a university intensive English program participated in the study. There were 15 students in the experimental group and 13 in the control group. All learners participated in three 50 minute game show activities. The experimental group and the control group received the same input and had the same opportunity to report noticing, but the control group seldom received interactional feedback. Learners' noticing was assessed through on-line learning journals, in which they made introspective comments while viewing classroom videotapes and questionnaire responses. Through a controlled pre- and post-test design, analyses of L2 development and noticing were carried out for each individual learner. Results showed that there was a positive relationship between noticing and interactional feedback. There was also a positive relationship between reports of noticing and development of one of the target forms (question forms).

More evidence about the beneficial role of feedback in promoting noticing and second language development can be found in studies that used a retrospective recall method (Kim & Han, 2007; Mackey, Gass, & McDonough, 2000; Mackey, Philp, Egi, Fuji, & Tatsumi, 2002), cued immediate recall (Philp, 2003), "on-line visually cued discrimination accuracy" (Trofimovich, Ammar, & Gatbonton, 2007), and a combination of a retrospective recall method and post-test scores (Egi, 2007b).

Russell and Spada (2006) synthesized recent findings obtained from both descriptive and experimental studies on oral and in written feedback, and concluded that corrective feedback is effective both in learners' oral and in written performance in general. Moreover, they found that the effect of corrective feedback is large and durable. This meta-analysis provides empirical evidence against Truscott's (1999) and Krashen's

(1994) argument that error correction is ineffective and even detrimental to second language development. Another more recent meta-analysis by Mackey and Goo (2007) reveals that interactional feedback[3] is one of the key beneficial features of interaction as measured by the short-term post-test scores, but not by delayed post-test scores. While empirical studies have demonstrated an overall effect of feedback, there is still much to understand about the relative efficacy of different types of feedback.

3.2 Empirical Studies on the Relative Efficacy of Feedback

If feedback is, in fact, facilitative in second language acquisition, the next question we seek to answer is whether certain types of feedback work better than others. Earlier work on corrective feedback, including Chaudron's study (1977) in French immersion classrooms, have already shown that not all feedback is noticed and incorporated by learners in the classroom. A challenge in determining which types of feedback are noticed is to devise a way of measuring the effectiveness of feedback. The descriptive studies reviewed in this section employed a number of measures on the efficacy of feedback including uptake and repair rate, conversational analysis, and private speech.

3.2.1 Descriptive Studies on Feedback

In their seminal work, Lyster and Ranta (1997) conducted an observational

[3] The term "interactional feedback" and "corrective feedback" are used interchangeably here. However, some scholars (e.g. Lyster & Mori, 2006) argues that "interactional feedback" is a more proper term because learners do not necessarily perceive all feedback types as corrective.

study of corrective feedback and learner uptake in four French immersion classrooms at the primary level. In their study, six types of feedback techniques were first identified and a model was developed to analyze transcripts of a total of 18.3 hours of classroom interaction taken from both subject-matter and French language arts lessons. Results include the distribution and frequency of the six different feedback types in addition to the distribution of different types of learner response following each feedback type. The findings indicate that teachers in French immersion classrooms tended to use recasts most of the time. In fact, recasts were used in over half of the total number of teacher corrective feedback turns (55%). However, they claim that recasts are the least effective in terms of eliciting learner repair in these French immersion classrooms. Four other types of feedback (elicitation, metalinguistic feedback, clarification request, and repetition) led to a greater number of student-generated repair moves. Based on this result, they hypothesize that the four latter types of feedback (i.e., negotiation of form) may engage learners more actively in a deeper level of processing by letting them retrieve the correct form themselves; recasts and explicit correction, whereas the former simply provide the correct form to them.

In his following articles, Lyster (1998a, 1998b) explains the differential effects of recasts and negotiation of form (which he later termed as "prompts") on eliciting immediate repair. In a further analysis of his data, he found that three quarters of teachers' recasts following ill-formed learner utterances were used in a similar way as non-corrective repetition after well-formed learner utterances. In the French immersion context, these identical functions of recasts and repetition may "override any corrective function that might have motivated the reformulations entailed in recasts" (Lyster, 1998b,

p. 188). The corrective potential of recasts may be further reduced by various signs of approval that teachers provide to confirm meaning. As a result, learners in such a context may perceive recasts as negotiation of meaning instead of negotiation of form. In other words, recasts may offer positive evidence in the same way as non-corrective repetition, but they may not be perceived by the students to offer negative evidence as many other researchers have claimed (Gass, 1997; Long, 1996).

In Canadian communicative ESL classrooms, Panova and Lyster (2002) found similar patterns of corrective feedback and learner uptake. A lower rate of uptake and repair followed recasts and a substantially higher rate followed negotiation of form. In contrast, Ellis, Basturkman and Loewen (2001) reported a higher level of learner uptake after the provision of recasts in their study when compared with the immersion and Canadian ESL contexts. Their data came from the observation of adult ESL communicative classrooms with a combination of form-focused instruction and meaning-focused activities in New Zealand.

Both the Panova and Lyster (2002) and the Ellis et al. (2001) observational studies used the same coding scheme of corrective feedback and uptake yet yielded different patterns of uptake and repair in relation to different types of feedback. The question that arises is: what factors may influence the rate of uptake and repair following different types of corrective feedback? A comparative study of these studies conducted in different instructional contexts may provide possible answers to this question.

3.2.1.1 The Issue of Context in Descriptive Feedback Studies

Sheen (2004) reviewed descriptive classroom feedback studies in four different contexts (ESL in New Zealand, ESL in Canada, French Immersion in Canada, EFL in Korea). She found that in more meaning or content oriented contexts, such as ESL in Canada (Panova & Lyster, 2002) and French immersion (Lyster & Ranta, 1997), recasts were less likely to be followed by repair and uptake than prompts, whereas in more form-oriented contexts, such as ESL in New Zealand (Ellis, Basturkman, & Loewen, 2001), or EFL in Korea (Sheen, 2004), recasts were equally effective at inviting uptake and repair. Suzuki (2004) confirmed Sheen's (2004) conclusion by investigating corrective feedback in a typical U.S. ESL context where people of many language backgrounds and nationalities were present in the same class. In her study, the uptake rate following recasts was much higher than Lyster & Ranta's (1997) study. Nicholas et al. (2001) investigated the relationship between recasts and instructional settings and concluded that:

> Taken together, the results of the classroom studies indicate that the context (particularly the communicative and/or content-based classroom) may make it difficult for learners to identify recasts as feedback on form and hence difficult for them to benefit from the reformulation that recasts offer. The exception may be some foreign language classrooms in which students' and teachers' focus is more consistently on the language itself. (p.744)

Lyster and Mori (2006) compared teacher-student interaction in two different instructional settings at the elementary school level (18.3 hours in French immersion and 14.8 hours in Japanese immersion in the U.S.). In their study, the immediate effects of explicit correction, recasts, and prompts (namely, rate of uptake following feedback) were investigated. The results showed a higher rate of student uptake and repair following recasts in Japanese immersion settings, whereas a larger proportion of repair

resulting from prompts was revealed in French immersion settings. Using the Communicative Orientation to Language Teaching coding scheme (COLT), (Spada & Fröhlich, 1995), Lyster and Mori identified Japanese immersion as characterized by an analytic orientation, which may have primed learners' attention to form.

These results lend support to the argument that the saliency and efficacy of feedback may vary across different instructional settings, highlighting the importance of distinguishing implicit and explicit feedback based on the discourse context (Ellis & Sheen, 2006). Research into the patterns and effects of corrective feedback in foreign language context in comparison with communicative or content-based second language classrooms may provide empirical evidence for this argument.

3.2.1.2 Descriptive Feedback Studies in Foreign Language Contexts

Despite some common characteristics in foreign language contexts, the way in which teachers in foreign language classrooms organize corrective feedback techniques as well as their pedagogical focus, may vary to a large extent. So far there have been only a few studies of feedback and learner uptake conducted in foreign language contexts, yet the mixed results of these studies render it difficult to draw a conclusion as to the effect of different types of feedback on learning a foreign language. This difficulty is largely due to the fact that these studies lack uniformity in design, both in the use of terminology and in providing a detailed description of the classroom context. My review of the following feedback studies, therefore, is based on the results of each individual study, with a discussion on the general trend at the end. The review begins with a look at how

foreign language classrooms are different from each other.

Some foreign language classrooms can be described as "analytic" while others may be characterized by "experiential teaching strategies" (Stern, 1990, 1992). The key differences between the two lie in the fact that the former strategy treats language as the subject of study and focuses on accuracy and error-free utterances, whereas the latter "focuses on content (subject matter, themes and topics of interest) rather than language *per se*", and therefore emphasizes fluency and meaning (Fazio & Lyster, 1998, p. 304), which resembles immersion or communicative ESL classrooms.

In order to study analytic language classrooms, Seedhouse (1997, 2004) adopted a Conversation Analysis methodology in analyzing classroom interaction excerpts and indicates that certain features of organization of repair are particular to this context. One such feature is that even when the learners produce utterances that are linguistically correct and appropriate, teachers might still repair them until the desired utterances are produced. Another very unusual phenomenon is that when a learner has failed to produce the targeted string of linguistic items, the teacher invites other learners to repair the learners' error, which is termed as "other-initiated other repair" (Seedhouse, 2004, p.147). Such kind of repair enables learners to focus on linguistic accuracy of the utterances without necessarily expressing their personal ideas and messages. In the meantime, this kind of repair also allows the students to become accustomed to the idea of peer-correction. Arguably, they will be able to help each other without hurting each other's feelings in the future (Seedhouse, 2004).

This kind of correction may also prime learners to develop a bias towards language form; therefore, a higher rate of uptake following a recast is expected in this

kind of context. In Lyster and Mori's (2006) comparative study, although Japanese immersion students are instructed in an overall content-based, communicative curriculum, some analytic teaching strategies were detected by the COLT scheme. Namely, these strategies were the use of choral repetition and an emphasis on speaking as skill practice in isolation through repetition and reading aloud. These analytic practices partially explained the larger proportion of uptake and repair following recasts (72% and 50%, respectively) in Japanese immersion than in French immersion (32% and 19%, respectively).

Similar findings were reported by Sheen (2004) in her observational study conducted in an Korean English as a foreign language context. In what she called "free talking" adult communicative classrooms, the uptake rate following recasts was 83%, a finding much higher than Lyster and Ranta's (1997) study in immersion contexts (31%). The repair rate of total uptake is also higher in Korean EFL contexts (70%) than in immersion contexts (57%). Sheen explained her findings in terms of the educational purpose and formal instruction which enabled learners to attend to the teachers' feedback and thereby notice the gap between their erroneous utterances and the teachers' correct form. Recasts in such contexts were more salient since they were often provided for a single linguistic error, which may also facilitate the opportunity for uptake after recasts, and thus explains the high uptake and repair rate.

The results of Sheen's study can be supported by another observational study in a foreign language context. In a German as a foreign language secondary school in Belgium, Lochtman (2002) conducted a descriptive study of corrective feedback. Tape-recordings of 12 lessons totaling 600 minutes were analyzed using Lyster and Ranta's

(1997) coding scheme. Interestingly, results in this study reveal that the majority (55.8%) of feedback types are prompts, which is different from previous studies in EFL and ESL contexts (Lyster & Ranta, 1997; Panova & Lyster, 2002; Sheen, 2004), where recasts are the predominant type of feedback. Furthermore, this study also reveals that recasts and explicit corrections elicit less but nonetheless successful uptake, whereas prompts result in more but less accurate uptake. The author found that "recasts in analytic FLT highly resemble explicit corrections and therefore appear to be fairly salient" (p. 276). In such a context where the focus of analytic teaching was on form, recasts not embedded in meaningful interaction elicited a higher uptake and repair rate (47.5% and 35%, respectively) than those in immersion contexts (31% and 18%, respectively).

Using a different research method, Ohta (2000) investigated the reaction to recasts by adult foreign language learners of Japanese. The focus of the study was on students' noticing of recasts directed to any member in the classroom. It was detected that students were able to respond to teachers' recasts in their "private speech"[4]. Furthermore, learners were more likely to react in private speech when it was directed at another learner or to the class as a whole, rather than when it was directed towards their own error. The results may have been influenced by several factors, however. As Nicholas, Lightbown, and Spada (2001) point out: first, the classroom had a strong focus on form, thus, the students were oriented to accuracy and language form rather than subject matter and other topics; second, the presence of individual microphones may make the students more aware of their language behavior. Finally, not all students wearing microphones were responding to recasts, which indicated that students' readiness of noticing and

[4] Private speech was defined as "oral language addressed by the student to himself or herself" (p.52). It was recorded by placing microphones on individual students during classroom interaction.

responding to feedback may vary even in the same classroom.

Overall, the studies of feedback in foreign language contexts present results that are more or less consistent in the following aspects: first of all, students in these contexts were able to react to feedback, irrespective of the feedback types, as measured either by uptake and repair (e.g. Sheen, 2004; Lyster & Mori, 2006) or private speech (Ohta, 2000). Second, it is likely that in form-oriented[5] foreign language contexts where students and teachers were mostly focusing on language form and accuracy, the corrective purpose of all feedback techniques seems more transparent. Furthermore, given that in foreign language settings, discourse tends to be repetitive and limited in the language used (Guthrie, 1984), short recasts with just one or two changes which are not embedded in meaningful interactions might be more easily recalled by learners (Philp, 2003) than recasts with multiple corrections combined with signals that confirm meaning.

Although descriptive classroom research on feedback provide an interesting yet complicated picture of patterns of feedback and uptake, many researchers (Mackey & Philp, 1998; McDonough, 2007; Philp, 2003) question whether there is a direct relationship between feedback and uptake, on the one hand, and between uptake and interlanguage development, on the other. For example, Mackey and Philp (1998) argue that uptake does not necessarily associate with interlanguage development and a lack of uptake or repair does not necessarily imply that learning is not taking place. As a result, experimental and quasi-experimental studies that employed a more rigorous pre-test, post-test design may provide more convincing evidence regarding these issues.

[5] The term "form-oriented" is used interchangeably with what Stern (1992) called "analytic" teaching context because they share similar characteristics such as the focus on language form and accuracy.

3.2.2 Experimental and Quasi-experimental Studies on Recasts and Prompts

Some observational studies (e.g. Lyster & Ranta, 1997; Panova & Lyster, 2002) have shown that in communicative or immersion contexts, recasts and prompts may have differential effects on eliciting immediate uptake and repair, which arguably indicate that these two types of feedback may have distinct functions in second language learning. However, the question still remains whether uptake is an indicator of noticing and a predictor for interlanguage development (Mackey & Philp, 1998; MacDonough, 2007). A review of experimental and quasi-experimental studies that directly measure the effect of learning through post-test scores may promote further understanding of the differential effects for recasts and prompts.

Motivated by both theoretical (i.e. the contributions of positive and negative evidence to L2 acquisition) and practical issues (i.e., what kind of form-focused instruction to recommend to teachers) (Ellis & Sheen, 2006), the effectiveness of recasts compared with other types of feedback has especially drawn researchers' attention during the past few years.

The distinction referred to hereafter draws on Lyster's (2004) classification of feedback; namely, *recasts* versus *prompts*. While the former type of feedback reformulates learners' erroneous utterances by providing the correct form, the latter provides various cues for learners to retrieve the target form themselves. As argued by Lyster (2004), although these four types of prompting moves (i.e. clarification request, repetitions, metalinguistic clues and elicitation) represent a wide range of feedback types, they all have one feature in common: they withhold correct forms and offer learners the

opportunity to modify their output themselves, whereas recasts provide learners with a covert reformulation. This implies that the two types of feedback moves may have distinctive functions in the acquisitional process: recasts provide positive evidence and exemplars to enhance connection in memory whereas prompts offer negative evidence that triggers retrieval from long term memory and restructuring of the interlanguage form. Before such a claim can be made, however, it is necessary to explore whether recasts are effective in second language learning. This is perhaps one of the most controversial issues in the feedback literature over the past few years (Ammar & Spada, 2006).

3.2.2.1 The Effect of Recasts in Second Language Learning

Some researchers have found that recasts (or negotiation of meaning) provide learners with implicit negative evidence and thus serve to benefit second language development in laboratory settings. For example, Mackey and Philp (1998) show that providing second language learners with intensive recasts, which focus consistently on one grammatical feature, is more effective for learners than interaction without recasts. In their study, 35 adult ESL learners from beginner and lower intermediate intensive English language classes participated in three sessions of communicative interaction (each lasting from 15 to 25 minutes) with native speakers. The recast groups received intensive recasts of their nontarget-like use of question forms from the interlocutor; the interactor group performed the same tasks but did not receive any form of feedback; and the control group only participated in the pre- and post-tests. Results suggest that learners at higher developmental levels who received recasts showed a greater increase in the

correct use of structures than learners who did not receive intensive recasts. It was also found that the presence of modified output after recasts did not seem to be an indicator of development of question formation. Therefore, they questioned whether learners' immediate responses to recasts are a good indicator of their subsequent use of recasts.

In another experimentally controlled condition, Long, Inagaki, and Ortega (1998) compared the effects of recasts and models on foreign language development of Japanese and Spanish. The treatment involved an information gap communication game. It was expected that the model and recast group would show greater development in the production of target forms than learners who did not receive any feedback. Positive results were found in the Spanish study in which recasts were more effective than models in the acquisition of adverb placement. However, the same effect was not found in Spanish object topicalization or in Japanese locative construction and adjective ordering. Long et al. explained their results in terms of structure difficulty and individual variations but nonetheless claimed that the results of the two experiments provided support for the facilitative role of negative feedback in second language learning.

The effects of recasts compared with models (or positive evidence) on the development of second language structures are rather unclear (Ishida, 2004). In some of the studies, the recast group performed better than the model groups (e.g. locative construction and adjective order rule in L2 Japanese (Mito, 1993)), but in other studies no significant differences were found in gains score for the recast group compared with the model group (Inagaki & Long, 1999). Ishida explains the mixed findings in terms of methodological discrepancies. She claims that in Inagaki and Long's (1999) study, the students in the model condition also had the output opportunity to repeat the model.

Ishida attributes this output opportunity to enhancing the salience of the positive evidence, which is equivalent to the recasting intervention that juxtaposes the interlanguage form with the corresponding target form. Another important factor that may affect the effectiveness of recasts is learners' prior knowledge of the target structure. Ishida believes that learners should have certain knowledge of the target structure in order to benefit from recasts. This issue has also been discussed by a number of scholars that take developmental readiness into account (e.g. Nicholas et al. 2001). Nicholas et al. commented that "recasts can be effective if the learner has already begun to use a particular linguistic feature" (p.752).

Following this line of thought, Ishida (2004) investigated the effects of intensive recasting in learning the Japanese aspectual form *–te i-(ru)* (a language feature for which learners already demonstrated partial knowledge), using a time-series design. Four college level learners participated in 8 conversational sessions. Overall accuracy increased in correlation with the number of recasts provided during the treatment sessions and, furthermore, the accuracy rate was retained in delayed post-tests. However, this study only had four participants and apparently no generalization is applicable to the research findings. Also, since there was no control group and the researcher employed a number of other feedback techniques (p.340) during the treatment sessions, it is questionable whether the increase in accuracy is only attributable to recasts.

In another small-scale study, Han (2002) investigated whether recasts would benefit the learning of linguistic forms that are already partially learned or in the process of being proceduralized (i.e., past tense in English). The subjects were eight adult female learners of English divided randomly into two groups (recast and control). They

participated in 11 sessions of written/oral narrative tasks with the researcher as the interlocutor over a period of 2 months. The researcher employed a pre-test, post-test and delayed post-test to measure the learners' tense consistency. Quantitative results of mean proportion disparity scores showed that the recast group developed a much higher control over tense consistency than the control group. Qualitative analyses of some learners' written narratives also revealed that recasts heightened the learners' awareness of tense consistency. Due to the small-scale nature of the study precluding the use of inferential statistics, it is difficult to evaluate Han's claim that this study provided "convincing evidence" (p. 565) on the positive effect of recasts on tense consistency.

In all of the aforementioned studies, the one-on-one nature of the interaction in performing those communicative tasks may have drawn learners' attention explicitly to the target feature, which Nicholas et al. (2001) claim to be a major concern in comparing experimental studies on feedback with classroom studies. As they suggest, the positive effects for recasts in the laboratory studies "may be due to the dyadic nature of the laboratory interactions, which may help learners recognize the interlocutor's feedback as corrective" (p.749). The few target structures which have been intensively recast may be easier for the learners to notice than recasts that were provided after a number of different grammatical features in classroom settings; therefore, the corrective nature of recasts is rather salient for the learners to recognize in laboratory settings.

Doughty and Varela's study (1998) was one of the first studies that investigated the effects of recasts in classroom settings. This study showed that "corrective recasting" preceded by a repetition of the learners' errors and emphasized with prosodic stress, was more effective than no feedback in learning simple and conditional past tenses for ESL

learners. The results of this study, however, need to be interpreted with caution (Ellis & Sheen, 2006; Lyster & Mori, 2006). The level of explicitness of "corrective recasting" and thus its effectiveness in this study is questionable. The "corrective recasting" in their study consisted of two phases: (1) "repetition to draw attention followed by (2) recasts to provide the contrastive second language forms" (pp.123-124). The operationalization of recasts in this way is in line with Lyster's (1998a) "repetition in combination with recast" (p. 68). In Canadian immersion contexts, uptake following corrective repetition in combination with other feedback types indicate that these combinations are relatively effective at eliciting repair and uptake (Lyster, 1998b). In effect, this type of recast already loses its implicitness as implied in its original definition by Long (1996), since repetition already draws learners' attention to the language form. The recast following repetition is more or less similar to explicit correction in providing positive evidence to the learners. As noted by Ellis and Sheen (2006), this type of didactic recast contains clear signals that make it explicit.

Another classroom study by Muranoi (2000) examined the impact of interaction enhancement on the learning of English articles. The researcher operationalized interaction enhancement as the teachers' provision of implicit feedback ("request for repetition" and "recasts"), together with debriefing on form or meaning. Ninety-one Japanese EFL learners participated in the study and were divided into three groups (Interaction enhancement plus formal debriefing group, Interaction enhancement plus meaning debriefing group, and Non-enhanced Interaction plus meaning focused debriefing group). The instructional treatment included three 30 minutes sessions of interaction in the target language. The researcher/teacher then provided feedback in

response to all errors involving indefinite article and some errors with tense-aspect forms to the treatment groups while the control group received feedback on meaning alone, and only when there was a communication problem. The groups then received debriefing either on form or meaning, depending on their pre-set conditions. Findings revealed that interaction enhancement had positive effects on the learning of English articles; and the group receiving interaction enhancement with debriefing on form outperformed the group receiving interaction enhancement with debriefing on meaning.

Although this study may provide evidence pertaining to the effect of implicit feedback on the acquisition of certain grammatical forms, it is difficult to tease the effect of implicit feedback apart from the effect of formal instruction (or "debriefing" in the author's term). Furthermore, similar to the recast group in Doughty and Varela's (1998) study, the treatment groups in Muranoi's study also received "repetition" together with "recast" as feedback. Therefore, it remains unclear whether it was the repetition or the recast that drew learners' attention to the target structure and contributed to subsequent development. More refined analysis of the effect of recasts without any intervening factors is certainly warranted.

3.2.2.2 Relative Efficacy of Recasts in Comparison with Other Feedback Types

The aforementioned experimental and quasi-experimental studies that either focused on recasts only or compared recasts with models or no feedback, irrespective of their methodological discrepancies and differences in terminology, showed some positive effect of recasts on the development of certain grammatical structures by second

language learners. Other recent studies that compared the effect of recasts and other types of feedback on second language acquisition yielded mixed findings. Some studies demonstrated the positive effect of prompts over recasts on the learning of certain grammatical features (Ammar & Spada, 2006; Ellis, Loewen, & Erlam, 2006; Ellis, 2007; Lyster, 2004b); while others did not show any differential effect of recasts and other types of feedback (e.g. Loewen & Nabei, 2007; McDonough, 2007). My review of these studies begins with classroom studies that compared the effects of recasts and other types of feedback, followed by studies carried out in laboratory contexts.

3.2.2.2.1 Classroom studies. Lyster's (2004b) study investigated the effects of form-focused instruction and corrective feedback on immersion students' acquisition of grammatical gender in French. Eight classes of 179 fifth-grade students participated in this study. Form-focused instructional treatments designed to draw learners' attention to selected noun endings that predict grammatical gender were implemented in the regular subject-matter instruction. During the 5-week instructional period, the three feedback groups also received different feedback treatment (ie., recasts, prompts, or no feedback), while the control group continued their normal classroom instruction. Based on the results of pre-tests, post-tests and delayed-post tests, each of which included two oral measures (object identification and picture description) and two written measures (binary choice and text-completion), Lyster found that form-focused instruction was more effective when combined with prompts than with recasts. Also, overall, form-focused instruction with feedback was better than no feedback as a means of enabling learners to

acquire French grammatical gender.

Ammar and Spada's (2006) quasi-experimental study investigated the potential benefits of recasts and prompts on the acquisition of a different structure: possessive determiners for French speakers in ESL contexts. Sixty-four students in three intact grade 6 intensive ESL classes were assigned to the recast group, prompt group and the control group. The treatment period was spread over a period of 4 weeks. Results revealed that all three groups benefited from the treatment, but the feedback groups showed superior performance than the control group. Furthermore, while the group receiving prompts significantly outperformed the recast group on written and oral post-tests, the effect of recasts depended on learners' proficiency levels. In particular, high-proficiency learners benefited equally from both prompts and recasts, whereas low-proficiency learners benefited more from prompts than from recasts.

Havranek (2002) carried out a quasi-experimental study on the relative effect of feedback on second language development in an Austrian EFL context. The study showed that the effects of various kinds of corrective feedback on second language development vary to a large extent. In this study, data were collected from 207 learners at six different age and proficiency levels, ranging from 10-year-old beginners to mature university students specializing in English. Using a tailor-made post-test design, the author presented results confirming that corrective feedback was effective in second language acquisition. Furthermore, findings showed that not only did the learner who initiated and engaged in feedback learn from these instances, but his or her peers (or auditors) who were present in the classroom also profited from feedback, more so when they were making silent responses and comparing their own hypothesis with the target

form provided by the teacher. In terms of the effect of different types of feedback, recasts without uptake (which was termed as "repetition" in this study) were the least effective, followed by recasts plus repetition. Elicited self-correction and other types of prompts resulted in significant improvement in the students' test scores.

As noted by Havranek (2002), there are certain conditions for the superiority of elicited self-correction (or prompts) over other types of corrective feedback: first, the learner's attention has to be drawn to the structure to be learned; second, the learner has to be actively involved in the interaction and must voluntarily make an effort to correct; and last, but not least, the learner must be developmentally ready for the structure to be corrected. This is also true for auditors who benefit from corrective feedback. Auditors have some advantages in incorporating feedback into their interlanguage in that they are freed from the high demand of on-line processing of classroom discourse; as a result, if they are ready for the target form, they have time to make the comparison of their own form with the teachers' target form. This could possibly explain why recasts, though less effective, still have an influence on learners' overall second language development of the corrected structure in foreign language classrooms. Similarly, this study also showed that corrective feedback addressed to one learner may have a potential facilitative role in other learners' second language acquisition. The limitation of this study, however, is that there is no pre-test that established baseline data in comparison with post-test scores, nor was there a control group; therefore, the findings could have been confounded by other factors than feedback alone.

Two recent classroom studies (Ellis, Loewen, & Erlam, 2006; Sheen, 2007) compared recasts with metalinguistic feedback. Both studies showed that the

metalinguistic group outperformed the recast group on the acquisition of the target feature (two functions of English articles in Sheen, 2007; and English regular past tense in Ellis *et al.*, 2006).

In Ellis *et al.* (2006), the effect of metalinguistic feedback and recasts on the acquisition of regular past tense in English *–ed* was examined. Participants were 34 low-intermediate level ESL students in a private language school in New Zealand. They were in three intact classes, which served as the three groups (the metalinguistic group, the recast group, and the control group). The two treatment groups received instructional treatment (two different half-hour communicative tasks), while the control group continued with their normal instruction. The relative effectiveness of both types of feedback was assessed by means of an oral elicited imitation test, a grammaticality judgment test, and a test of metalinguistic knowledge. Results showed that the explicit feedback (i.e., metalinguistic information) was overall more effective than implicit feedback (i.e., recasts). However, the effect was found mostly in the delayed posttest rather than the immediate posttest.

Sheen (2007) addresses the effect of different types of feedback on the acquisition of English articles and the extent to which individual differences mediate the effectiveness of corrective feedback. The study employed a quasi-experimental design in which 80 students in five intact classes participated. The participants were divided into three groups: the recast group, the metalinguistic group, and the control group. One difference between Ellis et al.'s (2006) study and Sheen's study was the operationalization of the metalinguistic feedback. In Ellis et al.'s study, metalinguistic feedback was similar to Lyster's (2004b) definition of prompts in that the teacher first

repeated the error and then supplied the metalinguistic information without providing the target form to the learner; in Sheen's study, however, metalinguistic correction was operationalized as teacher's provision of the correct form following an error together with metalinguistic information. Sheen made this distinction explicitly based on the argument that "the distinction between recasts and prompts is conflated with another important distinction, namely the implicit and explicit distinction" (p. 304). In doing so, she claimed that recasts and metalinguistic feedback in her study were both input-providing (Ellis, 2006), but different in terms of the degree of explicitness and the nature of the input provided in the feedback. Results showed that both in the immediate and delayed post-tests, the metalinguistic group outperformed the recast and the control groups. Recasts, however, did not show any significant positive effects.

The studies reviewed so far have compared recasts with one other type of feedback. The following study went a step further to compare recasts with both an implicit type of feedback that elicits self-repair (i.e. clarification request) and a more explicit type of feedback that elicits self-repair (i.e. metalinguistic feedback). Using a quasi-experimental design, Loewen and Nabei (2007) set out to investigate the following questions: (a) Does corrective feedback on English question formation errors during meaning-focused tasks lead to an increase in learners' performance on three types of tests that measure either implicit knowledge or explicit knowledge? (b) Is there a difference in the effectiveness of the three types of feedback for learners' performance on the three types of tests? Sixty Japanese EFL learners participated in the study. The recast group consisted of 10 students, while 8 students were in the clarification request group, 7 in the metalinguistic feedback group, and 10 in the no feedback group. The control group was

an intact class consisting of 31 students who received no treatment at all.

To answer the first research question regarding the effects of feedback on learners' performance on the three tests, the researchers found that the untimed grammatical judgment test and the oral production test did not show any increase in post-test scores while the timed grammaticality judgment test did. On the timed grammaticality judgment test, all the groups improved in their performance from pre- to post-test including the control group. Nevertheless, the feedback groups improved at a higher rate than the non-feedback groups. This result points to the fact that feedback may have an impact on learners' performance on a test that measures implicit knowledge. However, this effect was not shown in the other test (i.e. oral production test) that was also claimed to measure implicit knowledge. The authors explained this difference in terms of the receptive/productive nature of the tests and concluded that feedback may have an impact on the learners' ability to detect grammaticality of the sentences, yet may not impact their ability to produce them. An alternative explanation, however, might be that it takes a longer time to proceduralize the knowledge that they acquire and to make it accessible during oral production than to make on-line grammaticality judgments.

Another finding from this study that contrasts with previous studies (e.g. Lyster, 2004b; Ellis et al., 2006) is that no significant difference was found among the different feedback groups. Two possible explanations include the brevity of the treatment session (only 30 minutes) and differential amount of feedback provided among different groups (18 instances in the recast and elicitation groups, only 5 in the metalinguistic group) during the treatment session. The results should be interpreted with caution due to the small sample sizes in the treatment groups and lack of delayed post-test which may show

effects of feedback more effectively (Ellis et al., 2006).

3.2.2.2.2 Lab studies. Although classroom quasi-experimental studies generally demonstrated the overall beneficial effects of prompts (or metalinguistic feedback) over recasts, results obtained from laboratory studies are rather mixed. A study in a laboratory context was conducted by McDonough (2005), who investigated whether negative feedback and modified output produced in response to that feedback were significant predictors of development in second language learners' question formation, operationalized as stage development. Sixty Thai university EFL learners carried out a series of communicative tasks with native English speakers in four conditions that provided different feedback and modified output conditions. The "enhanced opportunity group" received repetition of the error with stress and rising intonation plus the opportunity to produce modified output; the "opportunity to modify" group received clarification requests plus opportunity for modified output; the "feedback without opportunity to modify" group only received repetition of the error with stress and rising intonation without the opportunity to modify output; while the "no feedback" group received neither feedback nor any opportunity to modify their output. The development of question formation was measured by an oral test after the treatment sessions.

Although statistical analysis did not show any significant differences between the feedback group and the no feedback group, the logistic regression analysis showed that modified output in response to feedback was the only significant predictor of question development. Furthermore, there were more learners in the "enhanced

opportunity to modify" group that produced more advanced question forms than in the "opportunity to modify" group. And learners in the "no opportunity to modify" group and the "no feedback" group did not produce stage 5 questions at all. This study thus provides empirical evidence for the output hypothesis and points to the importance of feedback and modified output in ESL question development.

Situated in the line of research that compares recasts with other types of feedback in lab contexts, McDonough (2007) carried out another empirical study that investigated the developmental outcomes associated with two types of interactional feedback, namely prompts and recasts. Different from the two studies mentioned earlier (Ellis et al., 2006; Sheen, 2007), this study did not compare recasts with a more explicit type of feedback such as metalinguistic feedback; instead, it compared recasts with clarification requests, which arguably also belong to the implicit type of feedback, yet differ from recasts in that they elicit responses more effectively. The study took place in a Thai EFL context and employed 74 participants. The participants carried out a series of communicative activities with native speakers within one week. Length of the treatment sessions, however, was not reported. Before and after the treatment sessions, they also participated in a pre-test and three post-tests measuring oral production tests. The measure of development in the study was operationalized as the emergence of new simple past activity verbs in all the three post-tests combined.

The results suggest that both clarification requests and recasts facilitated the emergence of simple past activity verbs. In addition, the findings revealed no advantage for clarification requests over recasts. Based on the analysis of audio-recordings of the treatment and testing sessions, McDonough (2007) also points out that clarification

requests are more effective in eliciting learners' responses than recasts, which confirmed previous findings. In the meantime, it was argued that lack of response after recasts may not necessarily indicate that learning did not take place. Although no significant differences were found between clarification requests and recasts on the development of simple past activity verbs, *post hoc* analysis did indicate that clarification requests facilitated the emergence of progressive activity verbs more than recasts. No significant difference between the recast group and the control group was found, however.

Notwithstanding its partially contradictory results from previous studies that showed differential effects between recasts and prompts (e.g. Lyster, 2004b; Ammar & Spada, 2006), one needs to bear in mind at least two factors that may have confounded the results of McDonough's (2007) study. The first one is that although the study was carried out in a school setting, the treatment sessions were conducted as one-to-one interaction. In such a context, the dyadic nature of the interaction excludes the possibility that recasts were interpreted as confirmation of meaning (Nicholas et al., 2001). In other words, recasts were likely to be as explicit and noticeable as other types of feedback, which can explain the fact that there was no significant difference between the recast group and the clarification request group in their development of simple past activity verbs.

The second factor is that development in this study was measured by emergence of new activity verbs, which followed the prediction from the line of research on tense-aspect hypothesis (Andersen & Shirai, 1996; Bardovi-Harlig, 1998, 1999, 2000). However, this way of measurement has rarely been used in the second language literature, as noted by the author herself. More longitudinal observational studies need to be

conducted to establish the validity of this type of measurement and the developmental sequences of simple past tense in English.

Another lab study by Lyster and Izquierdo (2010) sets out to investigate the differential effect of prompts and recasts on the acquisition of grammatical gender by adult second language learners of French. Twenty-five undergraduate students enrolled in an intermediate-level French course participated in the study. All students were exposed to a three hour form-focused instructional treatment on the target feature and each individual student participated in three different oral tasks in a dyadic interaction with a native or near-native speaker of French who provided either prompts or recasts in response to their errors. The study employed two oral production tasks and a computerized reaction-time binary-choice test. No significant differential effects between the recast group and the prompt group were found in this study. Instead, both groups significantly improved accuracy and reaction-time scores over time. It was argued that although no significant differences were found between the groups, these two types of feedback provide different opportunities for second language learning. Specifically, learners receiving prompts benefited from the repeated exposure to negative evidence as well as the opportunity to modify their output whereas recasts provided positive evidence as well as opportunities for learners to infer negative evidence due to the enhanced saliency of recasts in dyadic interactions.

The experimental and quasi-experimental studies reviewed here demonstrate mixed results regarding the relative efficacy of recasts compared with other feedback techniques. It is difficult to come to a firm conclusion regarding the relative efficacy of one type of feedback over the other (Ellis, 2007). What factors, then, may mediate the

effect of feedback? Do prompts have an overall more beneficial effect over recasts on all kinds of grammatical structures? Does linguistic structure dictate the type and effectiveness of feedback necessary? These questions are discussed in the following section.

3.2.2.3 Differential Effects in Relation to Grammatical Structures

After reviewing previous literature on the effect of different feedback techniques on second language acquisition, Ellis (2006, 2007) argued that a number of factors may account for the inconclusive nature of the relative effect of different types of feedback. These factors include different operationalizations of different types of feedback, the measurement of acquisition, and the choice of target grammatical structure. In terms of the choice of target structure, while some studies investigated morphological features (such as French gender in Lyster, 2004b; and possessive determiners in Ammar & Spada, 2006), others examined syntactical features (such as dative alternation in Carroll & Swain, 1993). As argued by Ellis (2007), developmental readiness as well as the complexity of the grammatical structure may reasonably be believed to influence the effect of feedback.

Based on the idea that feedback techniques may have differential effects on various target structures, Ellis (2007) carried out a study that involved two different grammatical structures at the same time. The two structures chosen were regular past tense "-ed" and comparative "-er". Based on a number of criteria including grammatical domain, input frequency, learnability, explicit knowledge, scope, reliability and formal

semantic redundancy, the two structures are hypothesized to differ in terms of grammatical difficulty. The research questions are: (a) Do recasts have a differential effect on the acquisition of the English past tense "-ed" and comparative "-er"? (b) Does metalinguistic feedback have a differential effect on the acquisition of English past tense and comparative? (c) To what extent does the effect of corrective feedback on the different grammatical structures differ according to type of feedback?

Three classes of students (n = 34) in a private language school participated in the study. The students in the treatment groups took part in communicative tasks which lasted for about an hour. The researcher provided corrective feedback either in the form of recasts or metalinguistic feedback to the students whenever they made an error in the target structure. In response to the first research question, the results showed no statistical difference on any of the scores for the two structures. In other words, the recast group did not show any significant gains over the control group on any of the measures. This result was different from other previous studies (Doughty & Varela, 1998; Han, 2002). Ellis explained the discrepancy of the results in terms of lack of saliency of recasts and short duration of the treatment.

In response to the second research question, overall, the metalinguistic feedback had a greater effect on the comparative. This could be explained by the fact that pre-test scores showed that learners in the study already demonstrated well-developed explicit knowledge of past-tense –ed but not of the comparative. As a result, the tests that were designed to measure explicit knowledge showed the greater effect of metalinguistic feedback on the comparative. Another interpretation would be that the well-established explicit knowledge on past-tense –ed may imply a ceiling effect for metalinguistic

feedback. In response to the third research question, differential effects were found only on the ungrammatical sentences of the oral imitation test. While the recast group did not show any significant difference on the two structures, the metalinguistic feedback group outperformed the control group on the ungrammatical sentences in the oral imitation test. However, no significant difference was found between the recast group and the metalinguistic group.

This study is the only one that investigated the effects of different types of feedback on different grammatical structures. It offers a new perspective in the plethora of existing feedback studies. However, just as Ellis (2007) concluded in his study, based on the paucity of studies that investigate the effect of feedback on different grammatical structures, what is needed in future research is to determine *how* linguistic factors may determine *when* different types of feedback will work for acquisition. Furthermore, as claimed by Ellis (2007), before enough evidence is established, it cannot be concluded that metalinguistic feedback is equally effective for all grammatical structures, nor can we generalize the effect of one type of feedback over another from studies that focused on only one structure.

3.3 Summary of Empirical Studies on Feedback

The studies conducted in both classrooms and laboratories that used post-test scores as indicators of feedback efficacy demonstrate, to a large extent, that corrective feedback is a "complex phenomenon with several functions" (Chaudron, 1988, p.152). In laboratory studies, recasts were shown to be more effective than "models" or no feedback

on learning certain grammatical features (Ishida, 2004; Han, 2002; Long *et al*., 1998; Mackey & Philp, 1998;); whereas in classroom settings, some studies reveal that recasts appeared to be less effective when compared to prompts (Ammar & Spada, 2006; Ellis *et al.*, 2006; Ellis, 2007; Lyster, 2004b).

As Nicholas, Lightbown, and Spada (2001) claim, the positive effects for recasts in the laboratory studies may be due to the consistent focus on a single structure and the enhanced saliency of recasts in dyadic interaction. The dyadic interaction draws learners' attentions to language forms more easily in laboratory than in communicatively oriented classrooms. Similarly, Mackey and Goo's (2007) meta-analysis confirms previous findings that laboratory and classroom settings provided different opportunities for development.

While context may influence the effect of feedback in both descriptive and experimental studies, it is only one of the many factors that make the issue of feedback so complex. Other factors include the use of different terminology and taxonomy of feedback (see a detailed discussion on this issue by Ellis and Sheen, 2006), amount of treatment, measure of development, (Ellis et al., 2006; Ellis, 2007), definition of acquisition (Han, 2002) and target structure (Ellis, 2007). For example, the amount of treatment can range from 30 minutes in Loewen and Nabei (2007) to 9 hours in Lyster (2004b) and 7-8 hours in Ammar and Spada (2006). Understandably, this wide range of treatment time can confound the results of these feedback studies. And this is only one example of the many different factors affecting the effect of feedback on second language learning.

Another important caveat to bear in mind is that none of the factors enumerated

are truly independent of one another (Mackey & Goo, 2007). Previous studies have shown that, for instance, the types of feedback provided, as well as number of uptake and repair instances, have been associated with contextual factors (Sheen, 2004), thus rendering it difficult to tease the effect of feedback types and that of context apart. Another example would be the link between types of linguistic structures and feedback type. Mackey et al. (2000) found that morphosyntactic features tended to elicit recasts while phonological and lexical features were more likely to elicit negotiation. In contrast, Lyster (1998b) found that in French immersion classrooms, teachers tended to recast grammatical and phonological errors and to negotiate lexical errors.

As the sheer number of feedback studies keeps increasing, a crucial step in future studies is perhaps to follow relatively established methodological procedures in feedback research while taking contextual factors into consideration in the design of the study. Building on the methodological framework of previous feedback studies, the rationale and hypothesis of the present study are outlined in the following section.

3.4 Rationale of the Study

While the association between contextual variables, feedback types, error types, and uptake has been established in previous studies, the extent to which feedback types may mediate different types of grammatical structures has rarely been studied. As Mackey and Goo (2007) correctly point out, due to the scarcity of empirical studies that compared recasts and metalinguistic feedback, any arguments for the efficacy of one kind of feedback over another are premature. Moreover, they point out the need for greater

theoretical specificity or practical motivations in making claims about the superiority of one feedback type over another.

Observational studies in classroom settings reveal that in form-focused foreign language contexts, recasts may be equally effective in eliciting uptake and repair (e.g., Sheen, 2004). Empirical studies that have compared the effects of these two types of feedback in laboratory and classroom settings have yielded mixed results, which have led to the speculation that recasts and prompts may have differential roles in channeling learners' attention to different aspects of grammar (Lyster & Mori, 2006; Ellis, 2007) and engage learners in different levels of processing (Lyster, 2004b).

My review of the literature in cognitive psychology and second language acquisition had led me to propose Skehan's (1998) dual-mode hypothesis as an alternative approach to understanding the role of different feedback types on different grammatical features. This hypothesis assumes that learners have at their disposal two kinds of learning mechanism, one being item-based, the other being system-based. Drawing on this line of argument, it is hypothesized that recasts and prompts may assist in the learning of different grammatical structure in different ways. In an attempt to test this hypothesis, the present study aims at investigating the differential effects of recasts and prompts on Chinese EFL learners' development of irregular past tense forms (item-based structure) and regular past tense forms (rule-based structure). The research questions and hypotheses are outlined in the following section.

3.5 Research Questions and Hypotheses

Based on the review of theoretical claims and empirical studies on feedback, it was hypothesized for the present study that the two feedback types (i.e. recasts and prompts) would have differential effects on the acquisition of rule-based versus exemplar-based structures. Specifically, recasts would favor the development of exemplar-based irregular past tense forms, while prompts would better facilitate the acquisition of rule-based regular past tense forms in English. The effects of corrective feedback on the learning of target forms were assessed by means of oral and written tests which are claimed to tap into both implicit and explicit second language knowledge (Ellis et al., 2006). The main objectives of my study were two-fold: (a) to investigate the overall effect of communicative activities combined with feedback on EFL learners' acquisition of English regular and irregular past tense, and (b) to compare the relative effect of two types of feedback on EFL learners' acquisition of regular and irregular past tense in English. Specifically, the questions that my study addressed are as follows:

Research Question 1

Q1: Do the groups that perform communicative activities while receiving feedback show an overall superiority in learning regular and irregular past tense over the group not receiving feedback[6]?

Research Question 2

Q2: To what extent do prompts and recasts differ in facilitating the acquisition

[6] Some studies (Ellis et al., 2006; Ellis, 2007; Sheen, 2007) include control groups that continue with their normal instruction without any kind of treatment. As a result, the beneficial effect of feedback group over the control group may be confounded by the communicative activities that highlight the target form. To investigate whether it is the sole effect of feedback that is beneficial for SLA, it is necessary to include a control group that also receive form-focused instruction, yet do not receive any feedback.

of regular English past tense?

Research Question 3

Q3: To what extent do recasts and prompts differ in assisting EFL learners' acquisition of irregular past tense?

Drawing on empirical and theoretical evidence reviewed in the previous chapters, three hypotheses were formulated to address the three research questions. Previous research showed that form-focused instruction, together with corrective feedback, is conducive to second language grammar development (Lyster, 2004b; Lyster, 2006). The overall effect of feedback has also been demonstrated by meta-analysis that showed beneficial effects of feedback in both oral and written format (Russell & Spada, 2006; Mackey & Goo, 2007). Furthermore, corrective feedback has been claimed to promote L2 learning because it prompts learners' noticing of L2 forms (Kim & Han, 2007; Mackey, Gass, & McDonough, 2000) and therefore assists in subsequent L2 development (Mackey, 2006; Tomosello & Herron, 1988, 1989). Hypothesis 1 was proposed with respect to research question 1.

Hypothesis 1: Both the prompt group and the recast group will outperform the control group on both oral and written measures at post-test and delayed post-test.

As predicted by Skehan's Dual mode system hypothesis, learners have at their disposal two kinds of systems: one being item-based and the other, rule-based. In addition, learners are constantly in search of external resources to modify their own

hypotheses about the target language (Izumi, 2003). Skehan (1998) also argues that the development of fluency depends, on the one hand, on the procedualization of the rule-based system with less control over the material, and on the other hand, on the memory-based chunks available for incorporation as wholes into language.

Previous empirical evidence suggests that prompts have an overall superiority in assisting the learning of rule-based grammatical structures (Ammar & Spada, 2006; Ellis et al., 2006; Ellis, 2007; Sheen, 2007) because they clearly provide negative evidence to the learner and signal that the learners' utterance is problematic. In addition, the self-repair process following prompts is claimed to help learners re-analyze what they have already learned at some level and restructure their interlanguage (Lyster, 2002). Furthermore, prompts may help learners gain greater control over already-acquired rule-based forms and access them in faster ways. Drawing on these theoretical claims, Hypothesis 2 is formulated as follows.

Hypothesis 2: Prompts will have more beneficial effects than recasts in learning regular past tense forms (rule-based structure) measured by both oral and written post- and delayed post-tests.

Evidence of classroom observational studies also revealed that when recasts are short, explicit, and involve only one or two changes, they are likely to be noticed by learners and may be conducive to learning as well. As Egi (2007a) claims, learners' selective attention to the negative evidence or positive evidence component of recasts differentially impacts L2 learning. In lexical learning, positive evidence may be more

likely to result in immediate interlanguage change than in morphosyntactic learning. The hypothesis to questions 3, is therefore based on the view that recasts will have a more beneficial effect in learning item-based items than prompts.

Hypothesis 3: Recasts, because they provide positive evidence, will better assist in learning irregular past tense forms (item-based structure) than regular past tense forms (rule-based structure) measured by both oral and written post- and delayed post-tests.

Chapter Summary

This chapter presents a review of previous studies on the effect of feedback in general as well as the relative efficacy of one type of feedback in comparison with other types of feedback. Drawing on the theoretical framework presented in Chapter 2 and results from previous feedback studies summarized in Chapter 3, the research questions and hypothesis of the present study were formulated. The next chapter will describe the research methods employed in the present study to test the research questions and hypotheses outlined in this section.

CHAPTER 4: METHODOLOGY

Despite a substantial amount of research conducted on the role of feedback on second language acquisition in general and the effect of specific types of feedback (such as recasts) on certain grammatical structures, so far there has been a paucity of research that unequivocally shows the beneficial role of one type of feedback over another on the acquisition of two different types of grammatical structures at the same time. Thus, the purpose of the present study is to improve on the research design of previous feedback studies and further probe this issue by investigating the effect of recasts and prompts on the acquisition of two types of grammatical structures (i.e. item-based versus rule-based structures) in a Chinese English as a Foreign Language (EFL) context.

This chapter begins with the design and procedure of the study, followed by a detailed description of the research context, participants, target structures, as well as the treatment and testing materials. Finally, it examines the coding and scoring procedures employed in the present study.

4.1 Research Design and Procedure

Previous classroom and laboratory studies of corrective feedback have yielded different results, since they either targeted patterns of feedback in classroom settings or the efficacy of feedback on the acquisition of one or two features in laboratory settings (Russell & Spada, 2006). The descriptive nature of many classroom studies restricted the scope of research to only examining the effect of feedback in the short-term (i.e., by

describing uptake rate only), whereas the laboratory studies failed to capture what would be the natural sequence of feedback and uptake in classroom settings, thus reducing the ecological validity of the research (Mackey & Gass, 2005). The present study overcomes these weaknesses by taking place in the classroom and thus being more ecologically valid, and by examining the effect of feedback over a long-term period rather than simply describing uptake rate.

4.1.1 Research Design

In order to investigate how EFL learners benefit from feedback in classroom settings in their acquisition of regular (rule-based features) and irregular (exemplar-based features) past-tense forms in English, the present study implemented a quasi-experimental design, with a view to closing the gap between classroom and experimental studies. Furthermore, the results from previous empirical studies on the relative efficacy of prompts and recasts in immersion and ESL settings (Ammar & Spada, 2006; Ellis et. al., 2006; Lyster, 2004b) may be validated by this study conducted in an EFL context.

Adopting a quasi-experimental design with three levels of between-subject variables (i.e., treatment conditions) and three levels of within-subject variables (i.e., testing time), this study sets out to compare the efficacy of two feedback types operationalized as "recasts" and "prompts" on the acquisition of two types of target structures: regular (-ed) and irregular past-tense forms in English. The study is quasi-experimental in the sense that the participants were not randomly assigned to different treatment conditions; instead, four intact classes composed of a total of 99 students

participated in the study. Care was taken, however, that the four intact classes were randomly assigned to the three feedback treatment groups: one class in the prompt group, one in the recast group and two in the control group. Students' acquisition of regular and irregular past-tense forms were assessed through one written production test and one oral production test before the treatment, and an immediate and a delayed post-test after the treatment. Figure 4.1 below illustrates the design and procedure of the research.

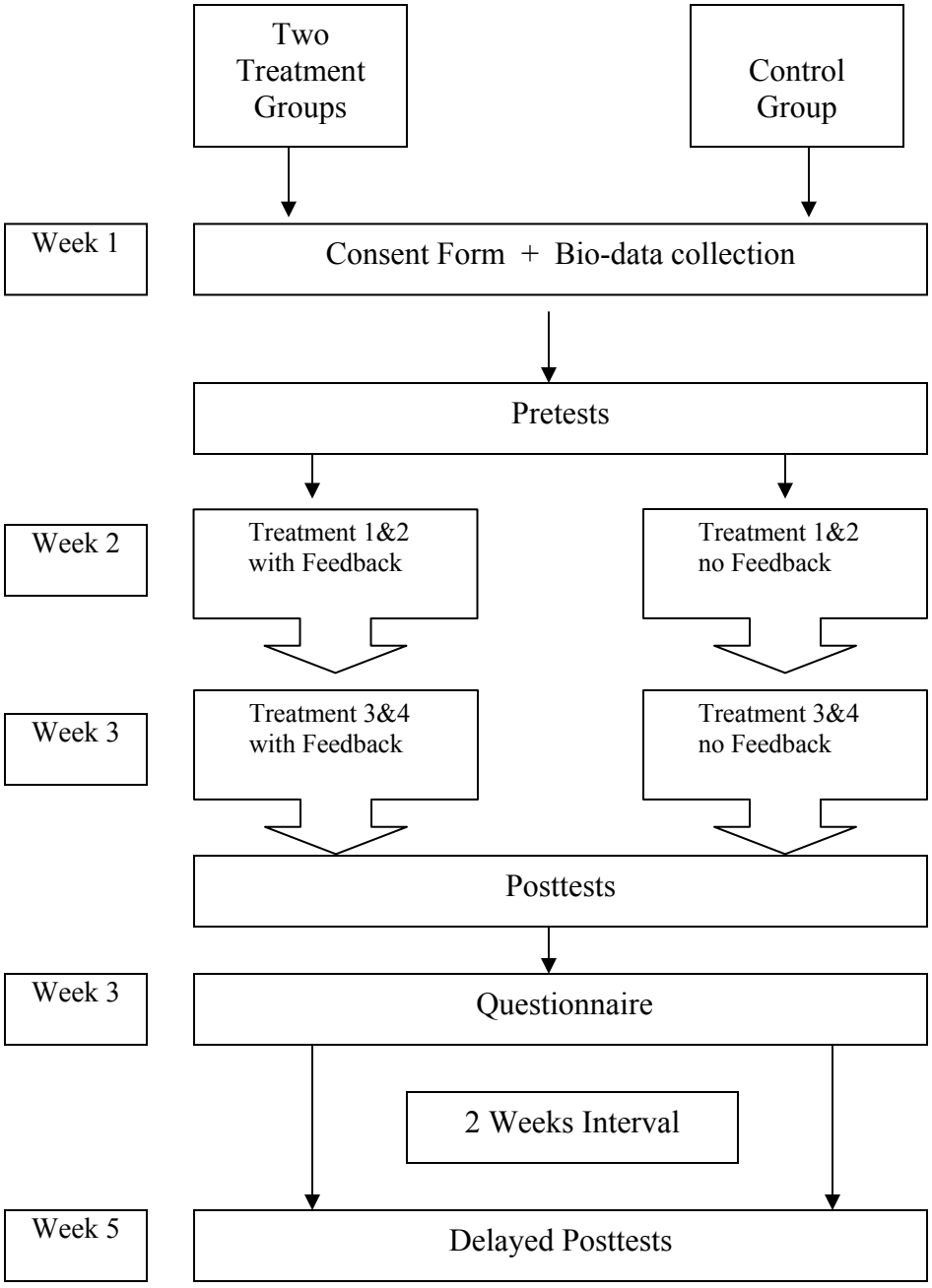


Figure 4.1. Design of the present study

4.1.2 Research Procedure

The research was carried out in university level EFL classrooms in Northern China in September 2007. Before the actual experiment, the researcher debriefed with participating teachers and provided them with the timeline of the study, general procedure, as well as treatment and testing material. The two teachers were also provided with a booklet specifying two feedback types with examples. Based on their own choice, the teachers and the researcher came to an agreement on what type of feedback they would provide during the treatment sessions. After the recruitment of participants, the researcher informed all participating students of the purpose and the procedure of the study. The collection of the bio-data questionnaire and consent form were administered before the actual treatment started. All the participants then took part in the pre-test.

During two consecutive weeks, participants in the treatment groups performed four communicative tasks, each lasting about 30 minutes, which were designed to elicit the use of regular and irregular past tense. The communicative tasks allowed students to practice the use of past-tense forms in meaningful contexts, since the function and use of past-tense forms had already been introduced at the secondary school English courses (Ministry of Education: New Standards in Secondary English Curriculum, 2003). In other words, students at this level already had the declarative knowledge of the target feature, yet had problems in using this knowledge under real operating conditions (Han, 2002). It is exactly at this point they needed practice that would allow them to incorporate meaning and form while maintaining the flow of communication. In the meantime, the teacher interacted with the students and provided feedback under designated conditions when errors occurred. In order not to draw learners' attention

exclusively to the use of past tense, teachers provided feedback on other kinds of errors as well, but they consistently provided one type of feedback only. The control group, in contrast, also performed the communicative activities, but did not receive feedback from the teacher during the interaction. In addition, the control group also participated in the three oral and written tests. The details pertaining to how the tests were administered to the control group will be discussed in the following sections.

Before and after the treatment sessions, with the purpose of tracking their development on the target form, participants were given a pre-test, an immediate post-test (immediately after feedback treatment) and a delayed post-test (two weeks later). These tests included both oral and written elements to test both the oral performance and grammatical awareness of the learners. A short exit questionnaire was also administered following the last treatment session to investigate students' overall focus during the treatment session as well as their attitudes towards the communicative activities and corrective feedback.

Data were analyzed using repeated measures ANOVA to find out (a) whether there was a significant effect of feedback on the acquisition of regular and irregular English past-tense forms, (b) whether feedback types influenced the way learners develop their proficiency in the use of regular and irregular past-tense forms in English, and (c) whether there was significant development of past-tense forms over time.

4.2 Research Context

The present study was carried out in an EFL context at the university level in Northeast China from the beginning of September, 2007, to mid-October, 2007. The participating university is a comprehensive university directly affiliated with the Chinese Ministry of Education. The study took place in the second year classes at the English department in the Faculty of Foreign Languages.

English is one of the three major programs in the Faculty of Foreign Languages. Each year, four classes in each grade level, each consisting of 20-30 students, are admitted to this program. These intact classes then undergo a four-year university education with a combination of diverse courses including mathematics, computer science, physical education and educational psychology. However, their curriculum mainly focuses on language training. This kind of training includes more than 20 hours of classes with a language focus each week, ranging from intensive reading, extensive reading, grammar and translation in their regular classrooms, to listening and audio-visual training courses in the audio-labs.

Students take a number of tests at the end of each term as well as two national exams for English majors throughout the country. The two national exams are called TEM-4 (acronym for Test for English Majors) and TEM-8 proficiency tests, in contrast with the CET-4 (acronym for College English Test) and CET-6 proficiency tests for non-English majors. These two national exams are designed to test English majors' comprehensive English abilities including listening, grammar, translation, writing, and dictation. TEM-4 is scheduled in April in the second semester of the second year in university while TEM-8 is scheduled in March in the second semester of the fourth year.

Each year, the Ministry of Education administers the date and time of the exams as well as the test content. All English majors are required to pass TEM-4, while passing TEM-8 is highly preferable but not obligatory.

4.2.1 Participating Teachers

The selection of teachers was based on their willingness to participate and their availability during the treatment sessions. Two teachers, together with the researcher, performed the treatment activities with the students during the two-week period. One of the participating teachers (Teacher A) was the teacher of the intensive reading course of two of the classes in second year. Due to the sick leave of the other teacher of the intensive reading course, Teacher A was also teaching intensive reading to the other two participating classes at the time of data collection. This led to a lack of teacher resources for performing the treatment activities concurrently in three different groups during the treatment period. Therefore, the teacher (Teacher B) who participated in the pilot study and the researcher were acting as substitute teachers for the recast group and the prompt group, while Teacher A conducted the activities with the two control group classes. All of the treatment sessions were carried out during the intensive reading class.

Both Teacher A and Teacher B were native speakers of Mandarin Chinese, and fluent speakers of English. In addition, they were both trained at the English Language and Literature program and received B.A. degrees from the same university, which is the university where the present study took place. Teacher A had just graduated from the English and American Literature master's program in this university and teacher B

graduated from the English Linguistics master's program in the same university 3 years ago. Neither of the two teachers had ever lived or studied in English speaking countries. Teacher B participated in the pilot study in July, 2007, and was therefore familiar with the procedures and material of the treatment. Since the pilot study was mainly aiming to test the treatment and testing instrument, she was not required to give feedback at that time.

The researcher, acting as Teacher C, was also a native speaker of Mandarin Chinese, and a fluent speaker of English. She graduated from another comprehensive university in the same city and received her B.A. in English Language and Literature and then an M.A. in English Linguistics. She is currently a Ph.D candidate majoring in second language education. She has two years part-time teaching experience and five years experience living abroad in an English-speaking country. The following chart outlines the bio-information and treatment condition of the participating teachers and the researcher who also participated in the treatment.

Table 4.1

Bio-information of participating teachers

Teacher Code	Gender	Age	Education	Experience in teaching	Treatment condition	Years abroad
Teacher A	Male	25	M.A.	1 year	Control	0
Teacher B	Female	29	M.A.	3 years	Recast	0
Teacher C (researcher)	Female	30	Ph.D candidate	2 years + Part-time	Prompt	5

The participating teachers chose their own feedback treatment preference after briefing with the researcher. Based on their choice, they were assigned to different treatment classes. The choice of treatment conditions for each participating class was completely random. Following a series of discussions of the general procedure and feedback sequence, participating teachers were also provided with a booklet with guidelines of feedback type. See Appendix A for the complete instructions provided to each participating teacher.

4.2.2 Participating Students

Prior to the current study, the researcher visited the site many times and piloted all the treatment and testing instruments, with the exception of the exit questionnaire, in a class at the same grade level as the participants. Based on the results of the pilot study, a

few instruments were revised or excluded[7]. The vocabulary level was also set at the appropriate level for the treatment and testing materials.

The recruitment of participants took place in September, 2007, following ethics approval from the McGill University Research Ethics Board. All 99 participants in the study were informed orally by the researcher of the general procedure as well as the purpose of the research and signed the consent form prior to the actual experiment (the ethics and consent forms are presented in Appendix A). All of the participants were in their second year of university, majoring in English language and literature.

Pre-tests were administered to select classes at the appropriate proficiency level (low-intermediate). Here proficiency level has a rather restricted scope, referring only to participants' proficiency level on the use of the target feature (Ammar & Spada, 2006). A majority of students (60%) in the present study started to learn English as a second language formally in school settings from Junior Middle School Grade 1 (at age 13 or 14 years old) and 40% of the students reported that they started to learn English in primary school or before primary school. The students reported an average of 7-13 years of experience in learning English, with an average of 9.25 years of experience.

Their ages ranged from 18 to 24, with an average age of 20 in each class. By this age, most students should have acquired a certain amount of vocabulary and basic grammatical structure to communicate in their second language, which enables the researcher to elicit and evaluate target structures in their oral and written tests. The students were uniformly from Chinese ethnic background, speaking Mandarin Chinese as their native language. There were 90 female students and 9 male students in the data set.

[7] Due to the overall high scores in the cloze test that was originally designed as part of the written tests, cloze tests were excluded from the real experiment. It is believed that the learners already developed declarative knowledge to score high (over 90%) in this type of test.

The class sizes ranged from 23 to 26 students. Out of these four intact classes, three groups were formed: one recast group (n = 25), one prompt group (n = 23), and one control group (n = 51). The questionnaire data about perceptions of activities and attitudes towards feedback included the complete data set (n = 94). Five students were missing from the questionnaire data collection session.

Table 4.2

Bio-data for participating students in the 4 classes

Classes	Treatment	No. of students	Male[8]	Female	Ave. Age	Ave. years of Eng. learning
1	Prompts	23	2	21	20	9
2	Control	25	2	21	20	11
3	Recasts	25	2	21	20	9
4	Control	26	3	23	20	8
	Total	99	9	90	20	9.25

4.2.3 Operationalizations

The two types of feedback were defined and exemplified as follows: Recasts were operationalized as a teacher's reformulation of a student's erroneous utterance, without changing the meaning of the student's original utterance, in the context

[8] The participating classes were all composed of English majors. It is a usual phenomenon that female students greatly outnumber male students majoring in arts subjects at the university level in China.

of a communicative activity (Sheen, 2007). The recasts could be full (see example 1) or partial — where the teacher only reformulated the incorrect segment (i.e. phrase, word) of the learner's utterance, as in example 2 below. The following examples were taken from Sheen (2007, p. 307).

Example 4.1: Recasts - Full Recast

S: There was fox.

T: There was **a** fox.

Example 4.2: Recasts - Partial Recast

S: He took snake back.

T: **The** snake.

Prompts were operationalized as one of four feedback types following Lyster and Mori's (2006) classifications: (a) metalinguistic clues, in which the teacher provides comments or questions related to the well-formedness of the student's utterance but does not provide the correct form of the target language (this is in line with Ellis et al.'s (2006) definition of metalinguistic feedback, but different from Sheen (2007) in which metalinguistic feedback includes the provision of the correct form); (b) repetitions, in which the teacher repeats the students' ill-formed utterances, adjusting intonation to highlight the error; (c) clarification requests, in which "an attempt was made to get learners to self-repair the erroneous utterance by asking for clarification" (Loewen & Nabei, 2007, p.367); (d) elicitation, in which the teacher directly elicits a reformulation from the students or pauses to allow the student to complete the teacher's utterance, or

asks the student to reformulate his or her utterance (Lyster & Mori, 2006). Examples of the three types of prompts are as follows. All of these examples were taken from the data of the present study.

Example 4.3: Prompts — Metalinguistic Clue

S: I went to the train station and pick up my aunt.

T: **Use past tense consistently**.

S: I went to the train station and picked up my aunt.

Example 4.4: Prompts — Repetition

S: Mrs. Jones travel a lot last year.

T: Mrs. Jones **travel** a lot last year?

S: Mrs. Jones travelled a lot last year.

Example 4.5: Prompts — Clarification Request

S: Why does he fly to Korea last year?

T: **Pardon?**

S: Why did he fly to Korea last year?

Example 4.6: Prompts — Elicitation

S: Once upon a time, there lives a poor girl named Cinderella.

T: **Once upon a time, there…**

S: there lived a girl.

4.2.4 Target Structure

Irregular and regular past-tense forms in English were chosen as the target structure for the current study with a view to comparing the relative effect of recasts and prompts in the acquisition of exemplar-based versus rule-based items in a second language. Another reason for choosing past tense as the target feature was that it is one of the features that is introduced early on in textbooks, yet difficult for learners even at intermediate or advanced levels to gain full control over (Ellis et al., 2006).

Regular English past-tense forms are considered to be a rule-based feature because there is a clear general rule (i.e. adding *–ed* at the end of the base form of a regular verb). Irregular English past-tense verb forms are considered to be exemplar-based or item-based (Ellis, 2005), because there is no such clear rule regarding the formation of the irregular past tense in English.

According to Quirk (1972), full verbs in English are classified as two morphological types: regular verbs (such as *call*) and irregular verbs (such as *drink*). For both types, the *–s* form and the *–ing* participle form are almost invariably predictable from the base form in the present tense.

It is in the past tense that these two types of verb forms differ. For regular verbs, if we know the base form (i.e. the dictionary entry form) of such a verb, we can predict the past form by adding *–ed*. This is a very powerful generalization because the vast majority of English verbs belong to this category. Irregular verbs differ from the regular verbs, however, in that the past form and the *–ed* participle of irregular verbs cannot be predicted by general rule from the base.

The past and the *–ed* participle forms of regular verbs have three pronunciations:

a) /id/ after bases ending in /d/ and /t/; e.g.:

pad ~ *padded* pat ~ *patted*

b) /d/ after bases ending in voiced sounds other than /d/ including vowels; e.g.:

buzz ~ *buzzed* budge ~ *budged*

call ~ *called* tow ~ *towed*

c) /t/ after bases ending in voiceless sounds other than /t/; e.g.:

pass ~ *passed* pack ~ *packed*

Irregular verbs differ from regular verbs in that either the past inflection, or the –ed participle inflection, or both of these, is irregular. The 250 or so irregular English verbs can be classified on the basis of certain criteria. Since it is impractical to account for both pronunciation and spelling together, only pronunciation will be considered in setting up classes of irregular verbs. Thus the criteria of classification are as follows:

a) Suffixation in V-*ed*$_1$ and/or V-*ed*$_2$, including not only the alveolar suffixes *–ed/-t* as in dreamed/dreamt, but also, for V-*ed*$_2$, nasal suffixes as in *shaken*, *torn*.

b) V –ed identity: i.e., V-*ed*$_1$ = V-*ed*$_2$, as in *met* ~ *met.*

c) Vowel identity, i.e., the various principal parts show no difference of base vowel: i.e., *put* ~ *put*

Table 4.3 shows how these criteria divide irregular full verbs into seven classes.

Table 4.3

Quirk's classification of Irregular verb classes (Quirk, 1972, p.111)

Class	Use of Suffix	V-ed Identity	Vowel Indentity	Example V	V-*ed*$_1$	V-*ed*$_2$
1	+	+	+	*burn*	*burned/burnt*	*burned/burnt*
2	+	+ −	+	*saw*	*sawed*	*sawed/sawn*
3	+	+	−	*bring*	*brought*	*brought*
4	+	−	−	*break*	*broke*	*broken*
5	−	+	+	*cut*	*cut*	*cut*
6	−	+	−	*strike*	*struck*	*struck*
7	−	−	−	*swim*	*swam*	*swum*

Characteristics of irregular verbs in the seven classes can be thus summarized as, for example:

Class 1: The suffix is used but voicing is variable (contrast *spent* with *made*). V-*ed*$_1$ and V-*ed*$_2$ are identical (*burned/burnt* ～ *burned/burnt*) and there is vowel identity in all parts (*build* ～ *built* ～ *built*).

Due to the highly complex and unpredictable characteristics of the simple past-tense forms of the irregular verbs and the highly rule-based past-tense forms of the regular verbs, the representation (Pinker & Prince, 1994; Pinker & Ullman, 2002) and acquisitional processes of the two kinds of verbs also presumably vary to a great extent (Ellis, 2005).

4.3 Treatment Instruments and Procedures

The following sections describe the treatment and testing material as well as the administration of these instruments in the classrooms.

4.3.1 Treatment Instruments

A number of criteria were considered in the selection of treatment tasks: (a) it must elicit the use of past-tense forms in learners' oral production; (b) it must be appropriate in the Chinese EFL classes in the sense that students need to feel comfortable communicating; (c) the focus on past-tense forms should be embedded in meaningful contexts so that the activities are not focus-on-formS[9] exclusively (Long & Robinson, 1998). In line with Swain's output hypothesis (1985, 1995), which claims that output production together with feedback may assist learners to compare their erroneous forms with the target form, four communicative tasks, all of which involve output production, were employed in the present study.

There were four treatment sessions, each lasting about 30 minutes. Each of these treatment sessions included one of the four communicative tasks developed by the researcher. In order to make sure that the students could actively participate in the communicative tasks, all of the four tasks had been piloted[10] in another class, as previously mentioned. These tasks were proven to be effective in eliciting the use of past-tense forms. In order to make sure that the vocabulary and content used in the tasks were

[9] Long and Robinson (1998) differentiate "focus on form" from "focus-on-formS" in that, in the former type of instruction, the primary focus is on meaning and communication with the learner's attention being drawn to language forms in context, whereas the second type of instruction resembles traditional grammar instruction which places a focus on forms in isolation.

[10] "A pilot study is an important means of assessing the feasibility and usefulness of the data collection methods and making any necessary revisions before they are used with the research participants" (Mackey & Gass, 2005, p.43)

commensurate with the students' level of proficiency, the researcher consulted with the teachers and referred to the students' textbook to tailor the material to their appropriate level. The four communicative tasks used in the study included two dictogloss tasks, one question-and-answer task, and one picture-cued narrative task. Figure 4.2 illustrates the sequence of the activities.

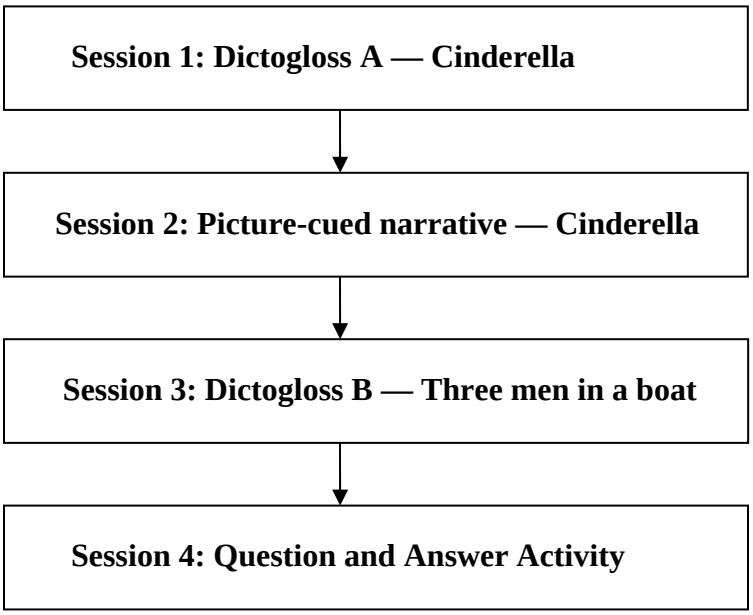


Figure 4.2. Sequence of treatment activities

4.3.1.1 Dictogloss Activity

The first type of communicative task used in these sessions was a dictogloss (Swain & Lapkin, 2001) for the purpose of raising learner's consciousness of past-tense forms. One of the dictogloss tasks was adapted from the famous fairy tale "Cinderella," the other from a humorous story, "Three men in a boat" written by Jerome K. Jerome. For a sample of the dictogloss activity, see Appendix B. Before the actual task started,

students were divided into pairs according to their seating positions. Altogether 10-13 pairs were formed in each class including the control group. After the instruction of the general procedure of the task, the teacher read the narrative story at a normal speed twice. In the first reading, the students were not allowed to take notes. In the second reading, the students took down what they heard from the teacher as much as possible. Since the text was read at a normal speed, the participants could not write down everything (Mackey & Gass, 2005). After the teacher read the whole story, the students in each pair compared their notes and reconstructed the text while maintaining the meaning of the original text. Then each pair was called on to narrate part of the complete story.

While the students were narrating the story, the teacher in the treatment groups provided either recasts or prompts to the learners' erroneous utterances. For the purpose of this study, the teacher mainly corrected past-tense errors while other errors (such as the use of vocabulary) were corrected when necessary so that the learners' attention was not drawn exclusively to past-tense forms. After the narration and correction phase, the teacher then provided learners with the original text with the past-tense forms highlighted in bold font. The students were asked to compare the text with their own with a view to further drawing their attention to the target form. In the control group, however, the teacher did not provide feedback while students were presenting the story to the class; instead, the teacher waited until they finished retelling the story and then handed out the original version of the story and asked them to compare it with their own text. The texts provided to the students were also highlighted in bold font as those provided to the other two groups.

4.3.1.2 Question and answer activity

In the second type of communicative activity, the students were given individual "question and answer" conversation practice with a focus on past-tense forms. Two sets of cards were prepared in advance by the researcher, one set for the students and the other set for the teacher. On each of the teacher's cards, there was a question about an action in the past. For example, *Did Mrs. Jones travel last year?* Below each question on the teacher's card, a specific noun or adverbial phrase (or clause) was presented (for example, *a lot*). On the corresponding student card, only the specific phrase "*a lot*" was given. If there were any verbs on the student card, they were all given in their base forms. After listening to the teacher's question, the students were expected to answer affirmatively first, and incorporate the information on their cards to complete the answer. For example, in response to the aforementioned question, the students were expected to say, "Yes, she/Mrs. Jones traveled a lot last year."

In this way, when the teacher asked the question, the students had to first listen to the question carefully, then answer promptly with the information provided. This process is claimed to tap into fast on-line processing (Skehan, 1998) and therefore elicit spontaneous speech data that may contain errors. Each student randomly selected one card from the pile, each with a number on top. The teacher had about 30 cards in hand, also with numbers on each card. When the session began, the teacher first called the number on top of the card and then asked the question written on the card. The student with the corresponding number stood up and first answered the question affirmatively,

followed by incorporating the information on their card to form a complete sentence as the answer to the teacher's question (See Example 4.7).

Example 4.7:

The teacher asked the question, "Did you lose your key?"

The student with the corresponding card with the information "on the way to school" would answer, "Yes, I lost my key on the way to school."

In answering questions about a particular third person referent, the students were allowed to replace the person's name with the corresponding pronoun, for example, "she" for "Mary" (See Example 4.8).

Example 4.8:

The teacher asked, "Did Mary lose weight?"

The student was expected to answer, "Yes, she lost a few pounds last week."

This kind of exercise aimed to elicit the use of regular and irregular past-tense forms, while still maintaining a communicative focus in the sense that it elicited spontaneous use of target forms in communicative contexts. More properly, perhaps, this kind of exercise resembled what DeKeyser (1998) terms as "communicative drills"[11] (pp. 52-53). This exercise was used in the pilot study and proved to be a very effective way to elicit past-tense errors from the students. It also allows the teacher to provide feedback in response to incorrect use of the target forms. In the treatment group, the teacher provided

[11] In contrast with "mechanical drills", communicative drills require the student to use the language to convey real meaning, while some rules that are taught previously, can be proceduralized in this process (DeKeyser, 1998, pp. 52-53).

feedback on the incorrect use of past-tense forms as well as other errors so that learners' attention was not drawn solely to past-tense forms.

In the control group, the teacher did not provide feedback; instead, the teacher continued the question and answer with the next learner. At the end of the activity, the teacher pointed out students' errors and wrote some common errors on the blackboard. The teacher also handed out a list of past-tense forms of all the regular and irregular verbs used in this activity to the students and asked them to refer to this list and reflect on their own mistakes. This kind of teaching method is in line with the traditional grammar-translation method and is still being used in a large number of classes in China (Hu, 2003; Yu, 2001).

4.3.1.3 Picture-cued Narrative Activity

The third type of activity used in the present study was a picture-cued narrative activity. This kind of activity has been used frequently in previous studies on past tense (Han, 2002; Ellis et al., 2006; Ellis, 2007).

In the general use of this activity, a set of pictures depicting a complete story that required the use of past-tense forms is presented to the students. The students are asked to individually narrate the story in sequence, using appropriate vocabulary and verb tenses. This kind of activity is more open-ended than other communicative activities, in the sense that no fixed vocabulary is provided or required. The advantage is that it motivates the students to use their second language creatively in a more authentic

communicative context, yet still remain focused on the proper use of tense and other aspects of grammar so that the story is coherent and accurate.

In the present study, in order not to overload the students with new information while they were communicating, pictures depicting the story of "Cinderella" were chosen. Since the students were already familiar with the storyline and some vocabulary from a previous activity (the dictogloss), they did not have to spend extra time searching for vocabulary or thinking about content. However, the pictures did not completely correspond with the story presented in the dictogloss.

In this task, the teacher first divided the class into small groups according to their seating positions, each group consisting of 2 or 3 members. Ten pictures, each with several characters depicting a scene in the story, were given to the pre-formed 10 groups. The students were required to prepare for 5 minutes, following which each student would present 2 sentences so that the whole class could tell a complete story. The students were also told that they could walk around the classroom to see other pictures to avoid repetition in content. When the students presented their part of the story, teachers in the treatment groups provided one type of feedback to students' errors, whereas in the control group, the teacher was only allowed to provide feedback on vocabulary errors. For a complete description and the administration procedures of these activities, please refer to Appendix B.

4.3.2 Treatment Procedures

Meetings with participating teachers were arranged two weeks before the corrective feedback treatments began. For the treatment groups, teachers were informed of the general purpose of the research objectives and procedures in advance. They were informed that the study was on feedback and interaction between teachers and students, but the hypothesis and the kind of feedback that each teacher was expected to provide was not disclosed. After a discussion with the researcher, the two teachers selected their preference for feedback technique according to their teaching practices and, based on their preference, a booklet with the timeline of the research, treatment and testing material as well as instructions for feedback was provided before the actual experiment. The following instructions with respect to feedback type were provided to the teachers.

4.3.2.1 Instructions for the Prompt Group Teacher

If the student makes an error in the use of past tense, you should use the following techniques to elicit the correct forms from the students, and allow them to say the correct forms themselves instead of giving them the correct forms.

For example,

Student: I make a cake yesterday.

Teacher should say

A: You make a cake yesterday?

Or

B: Do we say “make” for a past event?

C: What?

D: What tense do we use for a past event?

E: Yesterday, you…

The idea is that you withhold the correct form and use various cues to allow the students to correct their own errors.

4.3.2.2 Instructions for the Recast Group Teacher

If the student makes an error on the use of past tense, you should provide them with the correct form in a natural way so that the form becomes part of a correct utterance.

For example,

Student: I make a cake yesterday.

Teacher:

A: Oh, you made a cake. That's great. What else did you make?

or

B: You made a cake. Why did you make a cake?

C: Wonderful! What happened after you made the cake?

The key idea is that you correct the error, and at the same time, maintain the natural flow of the conversation.

4.3.2.3 Instructions for the Control Group Teacher

When the student makes an error in the communicative activity, you can temporarily ignore the errors and continue with the activities. At the end of the activity,

you provide the students with the list of regular and irregular past-tense forms and ask the students to reflect on them.

According to this information, the teachers were trained to provide either one type of feedback or no feedback to the students in the treatment sessions (the researcher was the teacher in the prompt group). The researcher and each teacher met 15 minutes prior to class and rehearsed how the teacher would provide feedback when the students performed the treatment tasks. The corrective feedback treatment took place in the four intact classes over a period of two weeks. Each treatment session was audio-recorded with a clip-on recorder attached to the teacher to ensure that the teachers consistently provided one type of feedback to the learners' errors. The analysis of the recordings of the treatment sessions in the three groups are presented in the results section in chapter 5. In order to measure students' development on the use of past-tense forms, two types of measurement were used.

4.3.3 Testing Instruments and Procedures

For each testing session (pretest, posttest and delayed posttest), an oral test and a written test were administered. Since learners in this particular form-oriented context were able to perform highly accurately on grammatical forms[12], tests were designed to measure learners' use of past tense forms in meaningful communicative context. In other words, the tests aimed to measure both the explicit and the implicit knowledge. According to Ellis (2004, 2005), the criteria for tests that measure implicit knowledge are

[12] As shown in the pilot study, learners in this context performed over 90% in the cloze test. This is partly due to their overall form-focused learning experience and therefore their predisposition to memorize the forms in decontextualized format.

(a) learners use their language by feel, (b) they have little need to draw on their metalinguistic knowledge, and (c) they are pressured to perform in real time with a focus on meaning. In contrast, tests of explicit knowledge need to elicit learners' use of rules, under no pressure, and encourage conscious focus on form and metalinguistic knowledge.

The oral narrative in the present study was designed to measure the learners' implicit knowledge (Erlam, 2006; Sheen, 2007) or spontaneous use of English past tense. The written narrative also allowed the use of past tense in a productive way, but differed from the oral narrative in that the learners had more time to reflect on their use of language. Therefore, the possibility that they may have drawn on both their explicit knowledge and implicit knowledge while performing the written task[13] cannot be ruled out. In addition, with the purpose of measuring the same construct as practiced during the practice sessions, both the oral and the written test included some past tense forms that appeared in the communicative activities during the treatment phase.

4.3.3.1 The Oral Test Session

The oral narrative task was adapted from an on-line grammar exercise (for reference, see Appendix C). It requires learners to retell a story based on a series of word cues. Oral narrative has been used either as treatment or test material in previous studies on past-tense forms (Ellis et al., 2006; Han, 2002). For this study, this test consisted of a short passage describing a crazy beach party in Korea and included 9 types (9 tokens) of regular and 17 types (18 tokens) of irregular past-tense forms in English. The students

[13] If the learners are consciously aware that the test is about the use of past tense, they may draw on their explicit knowledge. However, if they focus their attention only on writing a coherent story and focus on meaning, they may only draw on their implicit knowledge.

first read the story silently for three minutes; then they were presented with a series of word cues including prepositional phrases, verb phrases, and noun phrases to refresh their memory of the content of the story. In the meantime, learners were required to retell the story and to use the appropriate forms of the verbs provided as word cues. There was no planning time for learners after they saw the word cues.

All the verbs in the cues were in their base form; for example, "in the year 2000, fly…" Although there was no specific instructions on which tense should be used, the students were expected to retell the story in the past tense, given the adverbial phrase as a cue. For example, the student could say "In the year 2000, I flew to Korea for the first time," which is similar to the content in the original story. The test was time pressured to limit the learners' ability to draw on their explicit grammatical knowledge (Ellis 2005). Drawing on the results from the pilot study, all the content words that may have been difficult for students to recall (such as the name of the place or the person) were added in the final version of the word cues. This released the students' attention from retrieving the content of the story, since too much information load for the students may have deprived them of their ability to focus on form in the task (Skehan, 1998).

To control for test-retest effect, two versions of the oral test were used consecutively in the three tests; each version had exactly the same target regular and irregular verbs. However, the sequence of actions and people who performed the actions alternated in the two versions of the test so that the learners could not rely on their memory exclusively to perform the task. These two versions (Oral Test Form A and Oral Test Form B) were employed in the three testing sessions for all participants as A-B-A format. This also avoided participant fatigue that usually accompanies the repeated tests.

Since all the verbs in the two tests were the same, no reliability coefficient was calculated. The assumption was made that these two versions of the tests would reflect the use of past-tense forms equally effectively. The administration of the three tests is illustrated in the following figure.

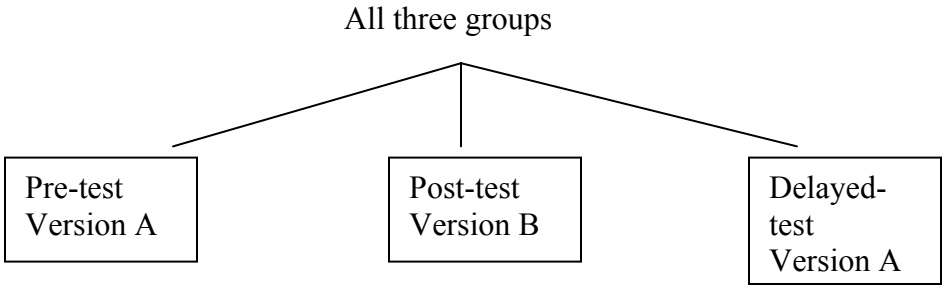


Figure 4.3. Sequence and versions of the oral tests

The three oral tests were conducted one week prior to the treatment, two days after the treatment sessions, and two weeks after the treatment sessions. All three tests took place in the listening class in an audio-visual lab in the department. For each testing session, each of the four classes participated in the tests consecutively based on their normal class schedule. The researcher and two research assistants administered the tests. The procedure of the testing session was as follows: after the students were all seated in a carrel with headsets with speakers on, the researcher provided test instructions to the students. The two research assistants then handed out the story and word prompts sheets and put them on each side of students' desk, with both pages facing down. Once the researcher said, "start", all of the students flipped the page with the story and read it through. After two minutes, the researcher required all the students to stop reading the

story, put the page down, and pick up the other page with word prompts and start retelling the story. In the mean time, the researcher started recording by pressing down the recording button on the control panel on the teacher's desk. After all the students finished retelling the story, the researcher and the two research assistants collected all the material and tapes.

4.3.3.2 The Written Test Session

The written test was designed in the form of a written narrative. The written narrative aimed to test the students' knowledge of the past-tense forms in production. The students were given the topic of the written narrative "A crazy day" on top of their test paper. They were required to use twelve verbs to describe a crazy day in 15-20 minutes. For the same reason as outlined in the previous section, there were also minor modifications of the written narrative tests for the three testing sessions. Specifically, three topics were provided in the three tests. In the first session, the topic was "a crazy day." The students could write anything that happened during a crazy day, using the twelve verbs. In the second session, the topic was to describe what happened on the tragic day that the Titanic sank. The topic for the third testing session was "Cinderella," a topic that the students were already familiar with from the treatment tasks. In order not to bring about item-learning for the students, the order of the 12 verbs in each testing session was altered randomly. In the third session, 2 new verbs were added to the list. An example of the instructions of the written narrative is given as follows:

Written Narrative (Post-test version)

Suppose you were a survivor of "Titanic." In twenty minutes, please use first person narrative to describe what happened that day. You should use all the verbs listed below and narrate the events in sequence. Use **SIMPLE PAST TENSE of these verbs ONLY** throughout the story (**Please do not use infinitive or participles or negation with these verbs**). However, you can add new verbs and use other tenses appropriately with the new verbs if necessary. Also try to connect sentences by using adverbs "first, then, consequently" and so on to write a coherent story.

Verbs to be used:

gather, smile, throw, mean, seem, draw, share, fear, flee, sigh, appear, sink

Twelve different verbs were provided in the test to measure learners' use of past-tense forms. Among these verbs, five were regular and seven were irregular. The choice of the verbs was based on the results of the written narrative from the pilot test. Verbs with which learners tended to make mistakes were selected as targets of the study. Care was taken that the meanings of these verbs were comprehensible to the students at this level. Furthermore, in order to ensure that the students used the simple past-tense form of these verbs, a note was added that forbade the use of infinitives, participles or negation of these 12 verbs. For example, according to the instruction, the students were not allowed to use "a party was thrown" or "the prince did not throw a party" because the involvement of the past participle, negation or the use of past continuous tense would not allow the researcher to capture the use of simple past-tense forms by the students[14]. The researcher also gave oral instruction to the students, emphasizing that all the words listed on the written sheet should be used as verbs[15]. The students were expected to describe the actions in sequence and use appropriate transitional words such as "after", "before", "finally" which were also listed on the paper. The target verbs remained the same

[14] Despite this effort, some students still used participle or negation of these verbs in their written narrative. The scoring procedure for these forms is discussed in the following section.
[15] In the pilot test, some words on the list were used as adjectives, (e.g., "mean") or nouns (e.g., "smile").

throughout the test, except the two new verbs added in the delayed post-test, but students were allowed to arrange the order of the actions freely and add details to make the story coherent.

The written test sessions took place immediately after each oral test session in the audio-visual lab. After the instructions were given by the researcher, the students started to write the narrative. After 20 minutes, the researcher and the research assistants collected the students' writings. In addition to the two measurements mentioned previously, a short questionnaire was administered in order to find out students' views about the treatment tasks and feedback.

4.3.3.3 Questionnaire

A short questionnaire with a combination of a "closed-item" and several "open-ended" items (Mackey & Gass, 2005, p.93) was administered immediately following the post test session. This questionnaire was adapted from Sheen's (2007) exit questionnaire designed to examine whether the students had become aware of the focus of the error correction treatments and tests (Ellis et al., 2006; Sheen, 2007). Two more open-ended questions were added to explore the learners' general attitude toward error correction and the communicative tasks used in the present study. The questionnaire was written in English, but the students were allowed to use Chinese to answer it so that they could provide details to the questions (For a complete version of the questionnaire, please refer to Appendix C). While the results of the questionnaire are discussed in the Results section, the following section offers a detailed description of the scoring procedures for each test.

4.4 The Scoring Procedure

The present study focused on the grammatical accuracy as well as the correct use of the target form in proper context with respect to regular and irregular past-tense forms. As a result, the learners' level of acquisition was measured in terms of how often these forms were supplied where they were required. This measurement/technique is known in second language research as "suppliance in obligatory context (SOC)" (Mackey & Gass, 2005, p. 232). Accuracy in the present study was operationalized as "the correct use of past-tense forms in appropriate past-tense context"[16].

The general criteria for coding and scoring the oral data were (a) the suppliance of the past-tense forms in obligatory context; (b) the accuracy of the past-tense forms of these English verbs. Specifically, the combination of (a) and (b) (i.e., the use of the correct form of the simple past tense of this particular verb in appropriate context) would grant the student "1" on the scoring chart next to the verb. For example, if the student said "I flew to Korea for the first time in the year 2000," next to the verb "fly" on the scoring sheet, the rater would mark "1" because he/she used the correct past-tense form of the verb "fly" (i.e., flew) in the appropriate past-tense context[17]. Similarly, as the student continued to narrate the story, the tense would already be set in the past (in the year 2000). Therefore, the student was supposed to use the past tense consistently throughout the text. The researcher consulted a native speaker of English and it was

[16] This is the researcher's own definition based on the type of tests in the present study. Since the written narrative is highly idiosyncratic, a universal obligatory context is not available. The raters therefore evaluated the contexts in each student's writing and the appropriate use of past tense forms in such contexts.

[17] In certain circumstances, the students would say a wrong form first and then self-correct. This would still be considered correct if the second form used was correct. Thus a mark of "1" would be given to that verb.

confirmed that all the target verbs should be used in simple past tense in the current context.

However, if the student failed to supply the correct form of the simple past tense or used other tenses (including other past-tense or present-tense forms), the scoring procedure would be as follows: a score of "0" would be marked next to the verb, if the student:

a) used other past tenses (e.g. past continuous tense, or past perfect tense, which would not allow the rater to ascertain whether the student knew how to use simple past tense of this verb properly.)
b) completely missed the verb (according to the context, the student used other verbs to replace this particular verb or simply ignored this part of the story and went on with narrating the next sentence). However, sometimes, the students would go back and narrate this sentence again later. In this case, the rater would still score the verb according to these criteria.
c) used the infinitive form of the verb (e.g. "We began **to jump** and run" instead of "We jumped and ran to keep warm") or the past participle (e.g. Ricky's hair was caught on fire" instead of "Ricky's hair caught on fire").
d) used the base form of the verb in past-tense context (e.g. "I fly to Korea in 2000").
e) used a wrong/hybrid past-tense form of this verb (e.g. I flied to Korea in 2000 or He was bleed badly).
f) used the present tense of the verb in obligatory past-tense context (e.g. It takes me 30 minutes to go back to Seoul the next morning.)

The same criteria were applied to both the oral and the written data. To score the oral data, the rater listened to the recordings of the students' performance and marked the accuracy of the past-tense form next to the corresponding verbs on the spreadsheet prepared as a coding sheet[18]. If the rater was not sure of the marking, she would listen several times until she could make a clear decision. To score the written data, the rater read through the students' writing samples and underlined the target verbs. Then, she went back to the text again and scrutinized the use of these past-tense forms, considering their accuracy as well as appropriateness within the context. In the meantime, she marked the use of past-tense forms on the spreadsheet with all the target verbs listed. The total of correct items became the participant's final score. In order to conduct statistical analysis, the total score of each participant was transformed to a percentage score with the formula: percentage score = total correct/total target items.

4.5 Inter-rater Reliability

As argued by Mackey (2005), regardless of the type of research and data coding methods, it is crucial to establish coding reliability. Because coding involves making decisions about how to classify or categorize particular pieces of data, it is necessary to employ more than one rater to increase the reliability of the research results. Care should also be taken to ensure that the second rater is carefully trained and kept relatively blind about which part of the data or for which group they are coding in order to reduce the possibility of coder biases.

[18] A spreadsheet with all the target verbs listed was used to code the data. See Appendix C for reference.

With these guidelines in mind, the researcher acted as the first rater and completed the entire coding and rating of 100% of the oral and written data. The second rater coded and rated 12% of the oral and written data. The second rater was a Chinese ESL teacher who received a master's degree from a Canadian University two years prior to the study. She had been teaching English in a Chinese middle school for over 5 years and had lived in Canada for over 5 years. She spoke Mandarin Chinese as her mother tongue and had native-like fluency of English.

In order to familiarize the second rater with the coding procedures, the researcher and the second rater met and went over the guidelines of the coding procedure together[19]. After the researcher made sure that the second rater understood all the procedures and instructions of the coding, the two raters together coded one student's oral and written data. The second rater was also provided with the coding sheet and was informed that if any questions arose, she could contact the researcher for clarification. A random selection of 12% of the oral and data from pre-test, post-test, and delayed-posttests of different groups was coded and rated by the second rater. The results of the coding by the two independent raters were compared using percentage agreement. The inter-rater reliability for the oral test data was 91.6% and 94.9% for the written test data. According to the guidelines provided by Portney and Watkins (1993), "for simple percentage, anything above 75% may be considered 'good', although percentages over 90% are ideal" (p. 244).

[19] The researcher provided specific details of how to code both oral and written data according to the "Scoring procedure" (Appendix C).

Chapter Summary

This chapter described the research methods employed in the present study. The study adopted a quasi-experimental mixed design with three levels of within-subject factor (testing time) and three levels of between-subject factor (treatment) to investigate the effect of recasts and prompts in comparison with no feedback in the acquisition of irregular and regular past-tense forms. In the four treatment sessions, participants performed four communicative activities with their teachers and received feedback or not according to designated conditions. In order to trace the participants' development in the use of the target structure, two testing measures including an oral narrative and a written narrative were employed.

The next chapter presents the analysis and results of the study with respect to classroom observations of feedback treatment, as well as repeated ANOVA results based on the participants' test scores across different groups over time and data from questionnaires.

CHAPTER 5: ANALYSIS AND RESULTS

This chapter presents the analysis and results of the data collected throughout the different stages of the study. It is divided into three subsections. In the first section, analysis of classroom transcripts of the treatment sessions are quantified and reported as baseline data that compare the quantity of errors, feedback and repair across different groups during the treatment sessions. The second section reports on learners' performance assessed by the oral and written tests on the past tense forms prior to, immediately after, and two weeks after the treatment sessions. Section three summarizes qualitative results from the questionnaires, reporting learners' major strategies in learning spoken English, attitudes towards feedback, as well as reflections on the communicative tasks and tests employed in the present study.

5.1 Analysis of Classroom Transcripts

This section presents analysis of classroom transcripts recorded during the treatment sessions across different groups. The analysis begins with coding categories and the procedures for transcribing classroom transcripts and then reports the results of the analysis. The purpose of presenting these results is to examine the implementation of different feedback treatment procedures during the treatment sessions and to compare the number of feedback and repair moves across different groups.

5.1.1 Coding Categories and Procedures

The treatment sessions were audio-recorded by a clip-on Sony digital audio-recorder attached to the teachers' clothing while the students were performing the communicative tasks and receiving feedback. The teachers walked around the classroom as the students were answering questions so that the students' utterances could be clearly recorded. The audio-recordings of the treatment sessions were then transcribed and analyzed by the researcher for the purpose of establishing baseline data with respect to the frequency of learner errors on past tense, the rate of feedback following learner errors, and the repair rate following teachers' feedback. This procedure was crucial in that it allowed the researcher to obtain a reference point in order to compare the number of errors and amount of feedback across different groups during the treatment sessions.

In analyzing the transcripts of the data, all simple past tense errors were identified and quantified as one of the following four possibilities: (a) failure to supply a simple past form in an obligatory context in a simple sentence (the use of the bare form or present tense form of the verb, as in Example 5.1); (b) overgeneralization of regular simple past tense form to an irregular verb (as in Example 5.2); (c) use of a hybrid form of auxiliary and simple past tense (as in Example 5.3); d) failure to maintain the use of the simple past tense form within a relative clause or compound sentence (as in Example 5.4).

Example 5.1

S: Once upon a time, there **live** a beautiful girl named Cinderella.

Example 5.2

S: The cat murmured “Meow”, which really **meaned** “cheer up”.

Example 5.3

S: Cinderella **must left** the party immediately.

Example 5.4

S: All the guests arrived before dinner **begin**.

All teachers’ responses immediately following learners’ past tense errors were examined and coded according to the definition and classification of feedback (recasts versus prompts) provided in Chapter 4. Only episodes that focused on past tense errors were considered and all feedback on meaning and other aspects of grammar were excluded from coding. Finally, the learners’ treatment session data were analyzed in terms of the learners’ responses to feedback. As shown in Example 5.5, a repair was identified as the learners’ provision of correct forms following the teachers’ prompts or recasts. In certain instances, learners’ repairs following teachers’ feedback were from fellow students instead of the learner to whom the feedback was directed. However, no distinction in the analysis was made in this respect. Therefore, the repair in the present study is a general term which encompasses self-generated repair as well as other-generated repair, unlike other researchers who differentiate between the two (such as Lyster & Ranta, 1997).

Example 5.5

S: All the guests arrived before dinner **begin**. (Error)

T: Before dinner what? (Feedback-prompt)

S: before dinner began. (Repair)

5.1.2 Results of the Analysis of Classroom Transcripts

The analysis of classroom transcripts begins with an overview of the distribution of errors, feedback, and repair across different activities. As Table 5.1 shows, dictogloss activities across all the three groups were less likely to elicit errors in the use simple past tense in comparison with the other two kinds of activities. In the two dictogloss activities, there were 16 and 14 past tense errors, whereas the other two activities elicited double these amounts. The question-and-answer activity had the advantage of eliciting both the learners' errors in the use of past tense as well as the number of feedback moves following these learner errors. A total of 37 errors occurred during this activity, followed by 25 instances of feedback. Eleven of the 25 instances of feedback were followed by repair. Overall, there was a total number of 108 errors in the use of simple past tense among all the students in the present study, 55 instances of feedback, and 27 instances of repair.

Table 5.1

Overall distribution of error, feedback moves, and repair across activities

Activity	Error	Feedback	Repair
Dictogloss 1	16	11	7
Dictogloss 2	14	9	5
Picture narrative	31	10	4
Question-answer	37	25	11
Total	108	55	27

Table 5.2 shows the distribution of errors, feedback and repair according to treatment group. As shown in the table, all three groups produced a relatively similar number of errors during the treatment sessions. In the control group, learners produced a total of 45 errors while in the prompt group and the recast group learners produced 31 and 32 errors, respectively. This is likely due to the treatment effect in the feedback groups in the first and second activities, which may have reduced the number of errors in subsequent treatment activities. Overall, the number of errors produced in each group was roughly comparable.

The number of instances of feedback following learner errors in the prompt group and the recast group were also comparable. Following 31 learners errors, the teacher in the prompt group provided a total of 27 prompts and 1 recast, whereas in the recast group, the teacher provided 23 recasts and 1 prompt following 32 errors on the use of past tense. Upon examination of Table 5.2, it can be seen that the teachers consistently provided the appropriate feedback type in their respective treatment group. Although this is to be expected in ideal situations, what tends to happen in natural classroom contexts is that the teachers mix different feedback types. The teachers in this study underwent prior training and rehearsal, however, and were therefore more consistent. The teacher in the control group also conformed to the rules of providing feedback only on content, with only three occasions of providing feedback on past tense errors. From the result of this analysis, it can be concluded that the treatment conditions generally conformed to the design of the present study.

In terms of the number of repairs following feedback in each group, in the prompt group, 26 out of 28 feedback moves were followed by repair, whereas in the recast group, there was only one repair and none in the control group. These results reinforce the argument that prompts create more opportunities for modified output whereas the nature of recasts obviates such an opportunity (Lyster, 2004b).

Table 5.2

Distribution of error, feedback moves, and repair across groups

Groups	Error	Feedback moves		Repair
		Prompt	Recast	
Prompt Group	31	27	1	26
Recast Group	32	1	23	1
Control Group	45	2	1	0

To summarize the results of analysis of classroom transcripts, the distribution of error, feedback, and repair across different activities revealed an unbalanced picture, with dictogloss activities eliciting fewer errors in comparison with the picture-narrative and question-and-answer activity. In addition, the question-and-answer activity was more likely to elicit errors as well as feedback following errors. In terms of the distribution of error, feedback, and repair across different treatment groups, all three groups produced a roughly comparable number of errors. The prompt group and the recast group received an almost equal number of feedback moves, yet the prompt group as expected outnumbered the recast group and the control group in terms of instances of repair following feedback.

5.2 Results of Analysis of Variance

This section focuses on the analysis of quantitative data on learners' use of simple past tense forms between groups across testing times (i.e., pre-test, post-test, and delayed post-test).

5.2.1 Data Set and Statistical Models

As mentioned in the methodology chapter, the three treatment groups in the present study vary in terms of group size. The prompt and the recast groups were relatively similar group in size, whereas the size of the control group was much larger because it included two classes that were taught concurrently by the same teacher at the time of the treatment. Since one of the assumptions of ANOVA is that group sizes are relatively equal, it renders the inclusion of two classes in the control group impossible. As a result, based on their bio-information, the classes that were relatively comparable in terms of duration of English learning were included in the statistical analysis, which resulted in one prompt group (n = 22), one recast group (n = 25), and one control group (n = 25). Altogether, 72 participants in three intact classes were included in the statistical analysis. However, the analysis of questionnaire data included all participants in the four classes (n = 99).

In analyzing the results of the study, two statistical models were used. The first model was a mixed design repeated measures ANOVA. This model was employed to determine (a) the differences in various treatment groups' use of simple past tense forms;

(b) learners' performance on the use of past tense forms across testing times; and (c) the interaction effect between treatment conditions and testing time (i.e., the differences in treatment effect across time) on learners' accuracy scores. Since post hoc multiple comparisons of pre-test scores revealed a significant difference between the groups on the use of irregular past tense forms in the written test, in addition to a repeated measures ANOVA, ANCOVA (Analysis of Co-variance) was used to determine any differences between groups at the time of post-test by adjusting for the pre-test scores. Results obtained from these analyses are presented in the following order: (a) results of the repeated measures of ANOVA of the test results across testing time; and (b) ANCOVA results on the use of irregular verbs in the written test.

5.2.2 Results of Mixed-model Repeated Measures ANOVA

Results from the pre-test, immediate post-test, and delayed post-test were analyzed using a mixed model repeated measures ANOVA. The results from the repeated measures ANOVA are presented in the following order: first, the results of repeated measures ANOVA on the use of overall past tense forms are provided, with the use of irregular and regular past tense forms in the oral tests grouped together with descriptive statistics as well as graphic presentations of the group means over time. Second, the ANOVA results on the use of total past tense forms, irregular and regular past tense forms in the written tests are presented together with descriptive statistics as well as graphic presentations of group means over time. Inferential statistics of post hoc multiple comparisons of group means are also displayed to reveal interaction effects between

treatment conditions and testing time. In all of the following analyses, the alpha level was set at 0.05.

5.2.2.1 Results of the Oral Test

This section presents results of repeated measures ANOVA on the use of overall past tense, irregular past tense forms and the use of regular past tense forms across the three groups over time in the oral tests.

5.2.2.1.1 Overall past tense forms. First, the descriptive statistics of the three groups' mean accuracy scores of the total past tense use as well as the standard deviations across the three testing times are displayed in Table 5.3. This is followed by Figure 5.1, which graphically presents the results of means of the three groups on the use of total past tense forms over time.

As shown in Table 5.3, all three groups improved their scores from pre-test to post-test and maintained this improvement to a varied extent in the delayed post-test. A repeated measures ANOVA confirmed that there was a highly significant effect of time factor $F(2, 61) = 19.02$, $p < .0001$. Post hoc comparisons of group means did not reveal any significant difference among the groups on the pre-test (see Table D2 in Appendix D for a complete pairwise comparison). Post hoc unconstrained (free combinations) step-down tests showed that all three groups significantly increased their scores from the pre-test to the post-test (for a table of pairwise comparisons of group means across testing time, see Table D1 in Appendix D). In addition, the prompt group made significant gains in its mean scores from the pre-test to the delayed post-test ($p < .05$). The control group

also significantly increased its mean scores from the pre-test to the delayed post-test ($p < .05$). ANOVA results revealed that there was no significant effect of treatment condition $F(2, 69) = 2.21$, $p = 0.11$, nor was there any significant interaction effect between group and time $F(4, 72) = 0.51$, $p = 0.73$.

Table 5.3

Group means and standard deviations on the use of overall past tense forms in the oral test

	Pretest		Posttest		Delayed Posttest	
Groups	*M*	*SD*	*M*	*SD*	*M*	*SD*
Prompt (*n* =22)	62.43	21.74	78.83	13.03	77.78	11.06
Recast (*n* =25)	64.27	20.49	75.60	19.56	68.41	25.09
Control (*n* =25)	53.70	19.73	69.07	14.93	64.44	16.09

Figure 5.1 plots the means of the three groups in the oral test over time and also shows that all the three groups increased their scores over time. It can be seen that the prompt group maintained the initial increase in scores on the delayed post-test better than the recast and the control group. However, post hoc multiple comparisons did not reveal any significant differences between the three groups at any point in time (see Table D2 in Appendix D for the complete pairwise comparison of group means at each testing time).

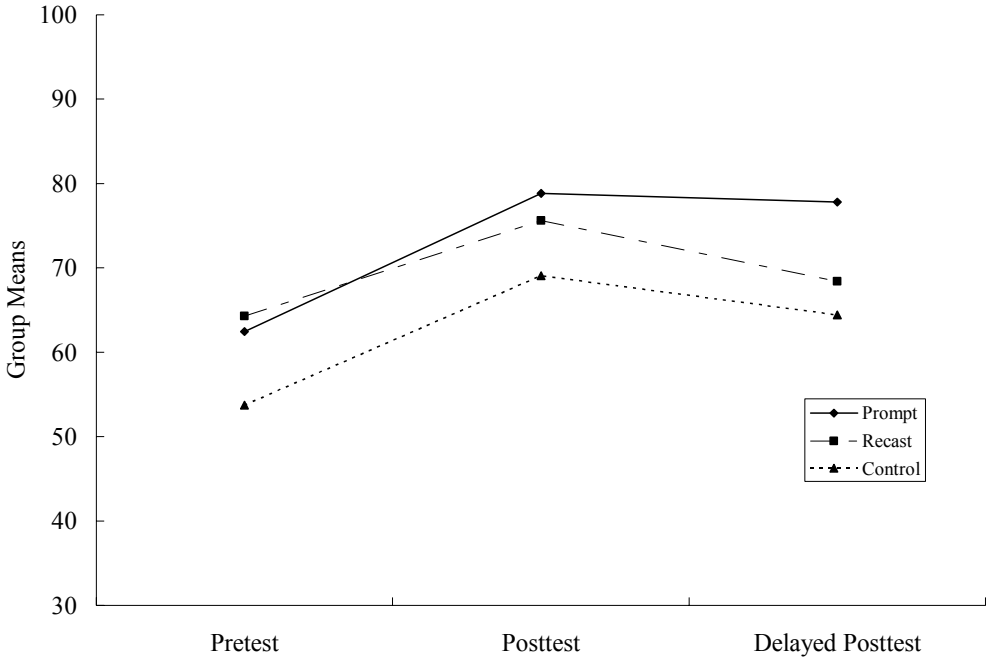


Figure 5.1. Group means on the overall use of past tense forms in the oral test

5.2.2.1.2 Irregular past tense forms. While the previous section presented results of the overall use of past tense forms, the present section reports ANOVA results as well as descriptive statistics related to the use of irregular past tense forms in the oral tests over time.

Table 5.4 displays the descriptive statistics regarding the learners' performance on the use of irregular past tense forms in the oral tests across time. As revealed in Table 5.10 and Figure 5.2, all three groups improved their accuracy rate in the use of irregular past tense forms from the pre-test to the post-test. The prompt group maintained the improvement in scores on the delayed post-test whereas the other two groups did not maintain the initial gain as well as the prompt group. A repeated measures ANOVA confirmed that there was a highly significant effect of time $F(2, 62) = 19.82$, $p < .0001$,

which means that, taken together, the three groups increased their scores over time. Post hoc unconstrained (free combinations) step-down tests did not reveal any significant difference among the three groups on the pre-test, but revealed that all three groups significantly increased their scores from the pre-test to the post-test (for a complete table of comparisons, see Table D5 in Appendix D). In addition, the prompt group demonstrated a significant gain in scores from the pre-test to the delayed post-test (p <.05). The significant gain in scores from the pre-test to the delayed post-test was also found in the control group ($p < .05$) (See Table D6 for a complete comparison of means). There was no significant effect of the treatment $F(2, 69) = 1.76$, $p = 0.18$), nor was there a significant interaction effect between group and time on learners' accuracy scores $F(4, 72) = 0.80$, $p = 0.53$.

Table 5.4

Group means and standard deviations on the use of irregular past tense in the oral test

	Pretest		Posttest		Delayed Posttest	
Groups	*M*	*SD*	*M*	*SD*	*M*	*SD*
Prompt (*n* =22)	63.64	23.99	78.95	13.30	76.14	11.91
Recast (*n* =25)	59.78	20.74	76.85	21.07	66.67	23.08
Control (*n* =25)	55.56	20.85	72.69	15.01	64.44	16.75

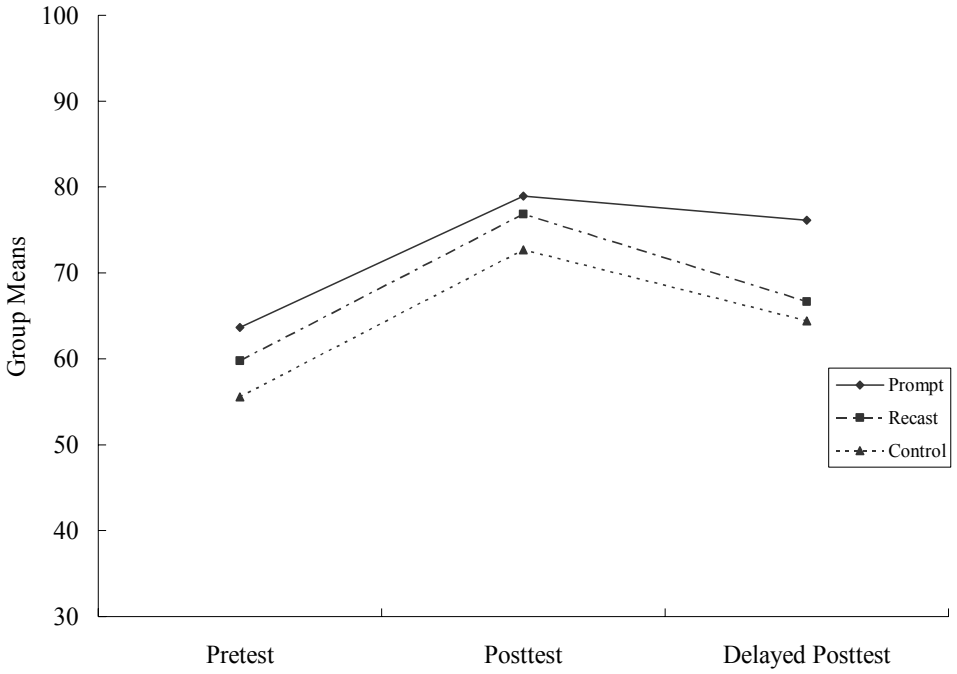


Figure 5.2. Group means on the use of irregular past tense in the oral test over time

5.2.2.1.3 Regular past tense. Table 5.5 shows the group means and standard deviations of the three groups' performance in the use of regular past tense across testing time. Similar to the irregular verb results, it can be seen that all three groups improved from the pre-test to the post-test. A repeated measures ANOVA confirmed that the effect of time was indeed significant $F(2, 61) = 8.40$, $p < .001$. However, there was no significant group effect $F(2, 66) = 2.87$, $p = 0.06$, or group and time interaction effect $F(2, 71) = 0.22$, $p = 0.92$. Overall, the three groups improved their scores at different points of time, as indicated by the group means. However, post hoc comparisons of multiple group means (see Table D7 in Appendix D for the complete table) failed to reveal any

significant difference within each group at different time points[20], nor was there any significant difference between groups at each time point (for a complete table, see Table D8 in Appendix D).

Table 5.5

Group means and standard deviations on the use of regular past tense in the oral test

	Pretest		Posttest		Delayed Posttest	
Groups	*M*	*SD*	*M*	*SD*	*M*	*SD*
Prompt (*n* = 22)	62.62	29.61	76.61	18.10	79.09	16.14
Recast (*n* =25)	61.33	24.03	72.22	17.57	70.56	22.30
Control (*n* =25)	54.67	22.89	64.82	18.44	64.44	20.58

Figure 5.3 graphically demonstrates the pattern of the three groups' performance over time. One interesting result is that, in contrast to the learners' performance in the overall past tense as well as irregular past tense use in the oral test, the three groups seemed to have improved their regular past tense scores on the post-test, and then maintained this improvement on the delayed post-test. Moreover, the prompt group not only maintained the increase in scores, but further increased its mean scores on the delayed post-test, although these increases in scores failed to achieve statistical significance.

[20] This conclusion is based on the adjusted *p* value for multiple comparisons. However, the raw *p* value did reveal significant differences between pre-test score and post-test score ($p < .05$) as well as between pre-test score and delayed post-test score ($p < .05$) for the prompt group (see Table D7 in Appendix D).

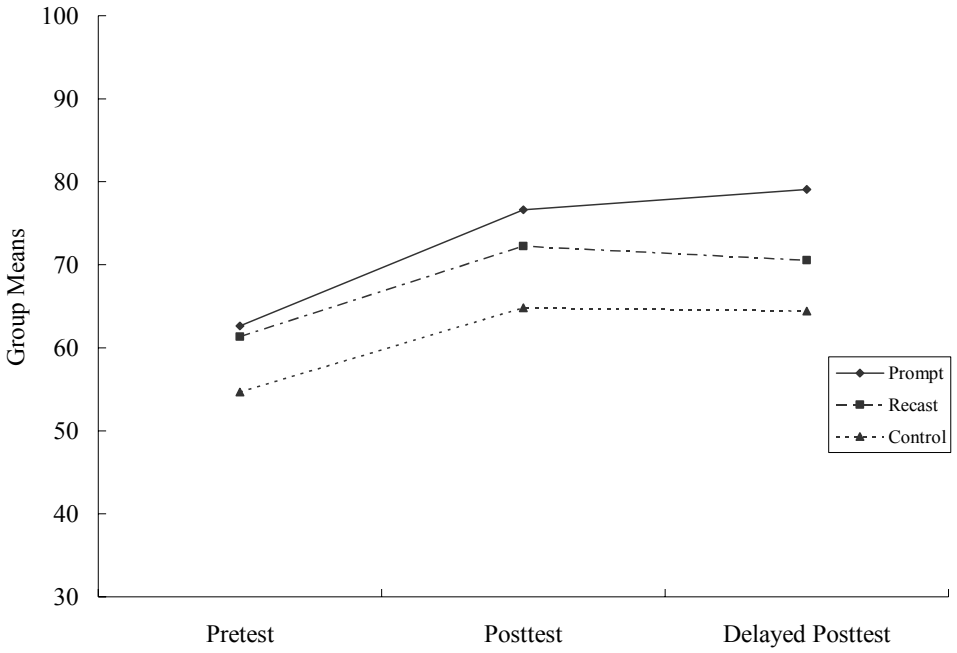


Figure 5.3. Group means on the use of regular past tense in the oral test

5.2.2.2 Results of the Written Test

While the previous section presented the results of the oral tests, this section focuses on the results of repeated measures ANOVA on accuracy scores of overall past tense forms, irregular past tense forms, and regular past tense forms across three groups over time in the written tests.

5.2.2.2.1 Overall past tense forms. First, the descriptive statistics of the three groups' mean accuracy scores of the overall past tense use appear in Table 5.12. This is followed by Figure 5.4, which graphically presents the results of means of the three groups' use of total past tense forms over time in the written tests. These results show that the prompt group outperformed the control group at the time of post-testing.

As revealed in Table 5.12, all three groups improved their scores from the pre-test to the post-test, as well as from the pre-test to the delayed post-test. Repeated measures ANOVA confirmed this result by revealing a highly significant time effect $F(2,67) = 44.96$, $p < .0001$. Post hoc unconstrained (free combinations) step-down tests that compared each group at different time points (see Table D3 in Appendix D for the complete table) show that both the recast group and the prompt group significantly improved their mean scores from the pre-test to the post-test at the .0001 level. In addition, the prompt group significantly increased its mean scores from the pre-test to the delayed post-test ($p <.0001$) and the recast group also achieved a significant gain in scores from the pre-test to the delayed post-test ($p <.05$). The control group, however, only significantly improved from the pre-test to the post-test ($p <.05$), but saw no significant difference in performance between the pre-test and the delayed post-test.

In addition to a significant time effect, a repeated measures ANOVA also detected a significant interaction effect between group and time $F(4,79) = 3.12$, $p < .05$, although there was no significant effect for group $F(2,68) = 1.47$, $p = 0.24$. Post hoc unconstrained (free combinations) step-down tests (see Table D4 in Appendix D for the complete table) that compare group means at each time point show that the prompt group outperformed the control group on the post-test ($p < .05$), and that the difference between the prompt group and the control group on the delayed post-test was approaching significance ($p = . 056$).

Table 5.6

Group means and standard deviations on the use of overall past tense forms in the written test

Groups	Pretest		Posttest		Delayed Posttest	
	M	*SD*	*M*	*SD*	*M*	*SD*
Prompt (*n* = 22)	54.55	20.21	87.12	12.77	80.68	14.86
Recast (*n* = 25)	60.87	16.75	81.52	13.52	76.52	15.78
Control (*n* =25)	61.86	19.17	75.00	17.35	68.33	16.84

Figure 5.4 plots the performance of the three groups at pre-test, post-test, and delayed post-test in the written test. As shown in this graph, the prompt group was the lowest in terms of group means on the pre-test; however, at the time of post-testing, the prompt group achieved the highest score, outperforming the recast and the control group. In addition, the prompt group maintained this advantage at the time of the delayed post-test. The recast group started at almost the same point as the control group, but also achieved higher scores on the post-test and the delayed post-test in comparison with the control group. However, the difference between the recast group and the control group was not statistically significant.

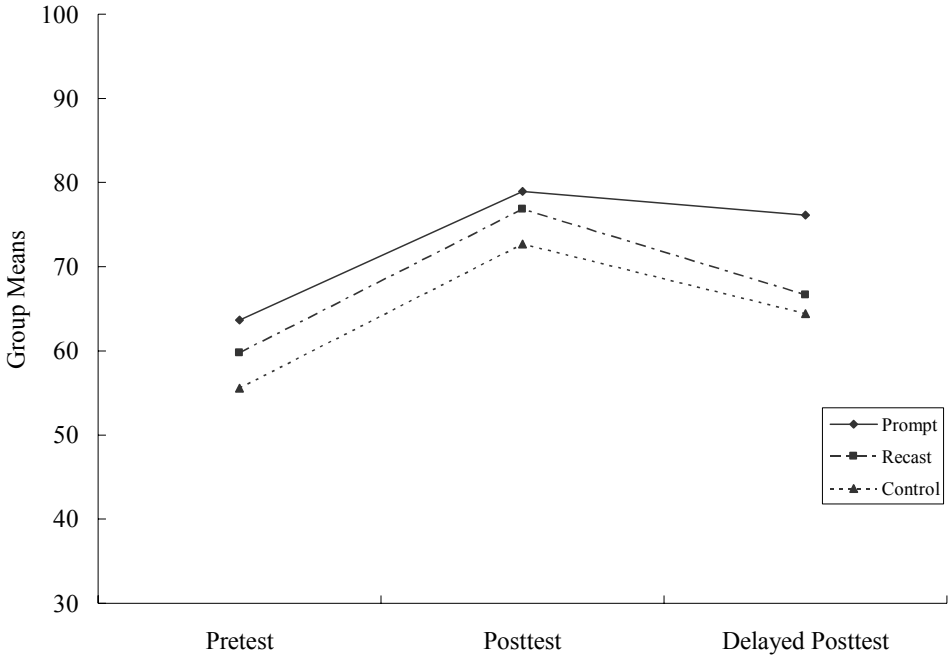


Figure 5.4. Group means on the use of overall past tense in the written test

5.2.2.2.2 Irregular past tense forms. Table 5.7 and Figure 5.5 together display the three groups' mean scores on the use of irregular past tense forms in the written test over time. The repeated measure ANOVA together with ANCOVA results show again that the prompt group outperformed the control group on the post-test.

A repeated measures ANOVA revealed a highly significant time effect $F(2,67) = 29.80$, $p < 0.0001$, as well as a significant group and time interaction effect $F(4,79) = 4.89$, $p < 0.05$. However, there was no significant group effect $F(2,68) = 0.68$, $p = 0.51$.

Table 5.7

Group means and standard deviations on the use of irregular past tense in the written test

Groups	Pretest		Posttest		Delayed Posttest	
	M	*SD*	*M*	*SD*	*M*	*SD*
Prompt (n = 22)	39.09	23.48	77.27	27.11	75.45	20.41
Recast (n = 25)	42.61	25.08	67.83	25.40	65.45	26.32
Control (n = 25)	53.85	23.85	60.80	26.13	59.20	21.97

Since post hoc unconstrained (free combinations) step-down tests (see Table D10 in Appendix D for a complete comparison) that compare groups at each time point detected a possible significant difference between the prompt and the control group[21] in terms of the raw scores (p = 0.036), an ANCOVA was employed for further analysis of the post-test scores by adjusting for the initial discrepancies on the pre-test.

The ANCOVA results demonstrated that there was a significant difference among the three groups at the time of immediate post-testing, after adjustments for differences on the pre-test $F(2, 68) = 3.83$, $p < 0.05$. The post hoc comparisons, using the Tukey-Kramer adjustment for multiple comparisons, revealed a significant difference between the prompt group and the control group ($p < 0.05$). This result further confirmed the repeated measures ANOVA result which had revealed a significant group and time interaction effect on the use of irregular past tense forms in the written test.

[21] The tests showed that the difference between raw scores was significant (p <.05) but the p value adjusted for multiple comparisons was not (p = 0.21). Given that there is a possibility that the differences in the post-tests may have been influenced by this initial difference in the pre-test, an ANCOVA with pre-test score as a co-variate was employed to further analyze post-test scores.

In addition to the ANCOVA result, post hoc unconstrained (free combinations) step-down tests (see Table D9 in Appendix D for a complete comparison) that compared each group at different time points confirmed the significant interaction effect between group and time. The prompt group significantly improved from the pre-test to the post-test (p <.0001) and maintained this improvement in scores at the time of the delayed post-test (p <.0001). Similar results were found for the recast group which also demonstrated significant gains from the pre-test to the post-test (p < .001) as well as from the pre-test to the delayed post-test (p < .001). However, the control group did not improve significantly from the pre-test to the post-test (p = 0.63), nor from the pre-test to the delayed post-test (p = 0.77).

Figure 5.5 shows the pattern of the three groups' performance over time in the use of irregular past tense forms. The graph shows that the prompt group outperformed both the recast and the control group at the time of the post-test and the delayed post-test, although it started with the lowest mean score at pre-test. The recast group also increased its mean scores from the pre-test to the post-test and maintained this increase in scores at the time of the delayed post-test. However, the control group improved less substantially in comparison with the feedback groups, which is demonstrated by the flat line.

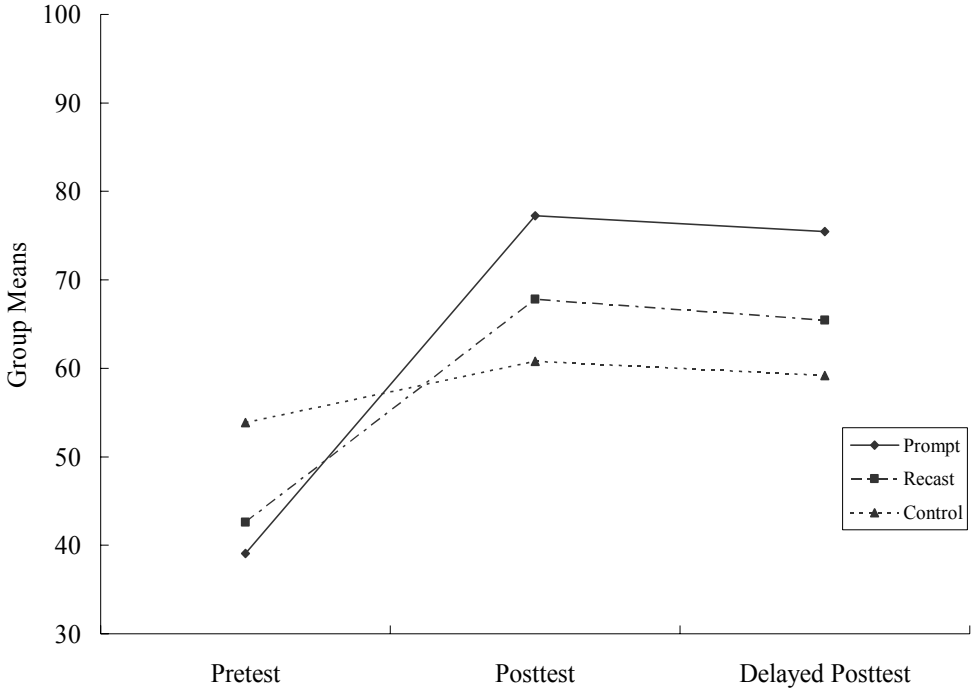


Figure 5.5. Group means on the use of irregular past tense in the written test

5.2.2.2.3 Regular past tense forms. While the previous section focused on the use of irregular past tense forms over time, this section presents the results of three groups' use of regular past tense forms in the written test over time.

Table 5.14 displays the group means and standard deviations on the pre-test, the post-test, and the delayed post-test. In comparison with their mean scores in the use of irregular past tense forms, the three groups performed better on their use of regular past tense forms at the time of the pre-test, all three groups achieving over 60%. A repeated measures ANOVA confirmed the findings in the descriptive statistics, revealing a highly significant time effect $F(2,67) = 31.88$, $p < 0.0001$. However, ANOVA did not detect any significant group effect $F(2,69) = 2.78$, $p = 0.07$, nor was there a significant interaction effect between group and time $F(4,80) = 0.94$, $p = 0.44$. Post hoc unconstrained (free

combinations) step-down tests did not detect any significant difference among the groups at the time of the pre-test, yet they show that the three groups significantly improved their scores from the pre-test to the post-test. In addition, only the prompt group maintained the initial gain in scores at the time of the delayed post-test, as demonstrated by a significant difference between pre-test scores and delayed post-test scores ($p < .05$) (see Table D11 in Appendix D for the complete table). However, post hoc comparisons of group means at each time point did not reveal any significant difference among the groups (see Table D12 in Appendix D for the complete table).

Table 5.8

Group means and standard deviations on the use of regular past tense in the written test

	Pretest		Posttest		Delayed Posttest	
Groups	*M*	*SD*	*M*	*SD*	*M*	*SD*
Prompt (*n* = 22)	65.58	21.89	94.16	8.43	84.42	15.84
Recast (*n* = 25)	73.91	19.09	91.30	11.18	84.42	13.88
Control (*n* = 25)	67.58	26.36	85.14	14.57	74.88	17.14

Figure 5.6 graphically presents the three groups' performance in their use of regular past tense in the written test over time. It can be seen that the prompt group started with the lowest mean score at the time of the pre-test, but achieved the highest mean score on the post-test. The recast group and the control group showed almost parallel increase from the pre-test to the post-test, but the recast group maintained its gains better than the control group at the time of the delayed post-test. However, these results failed to achieve any significant difference among the groups at each time point.

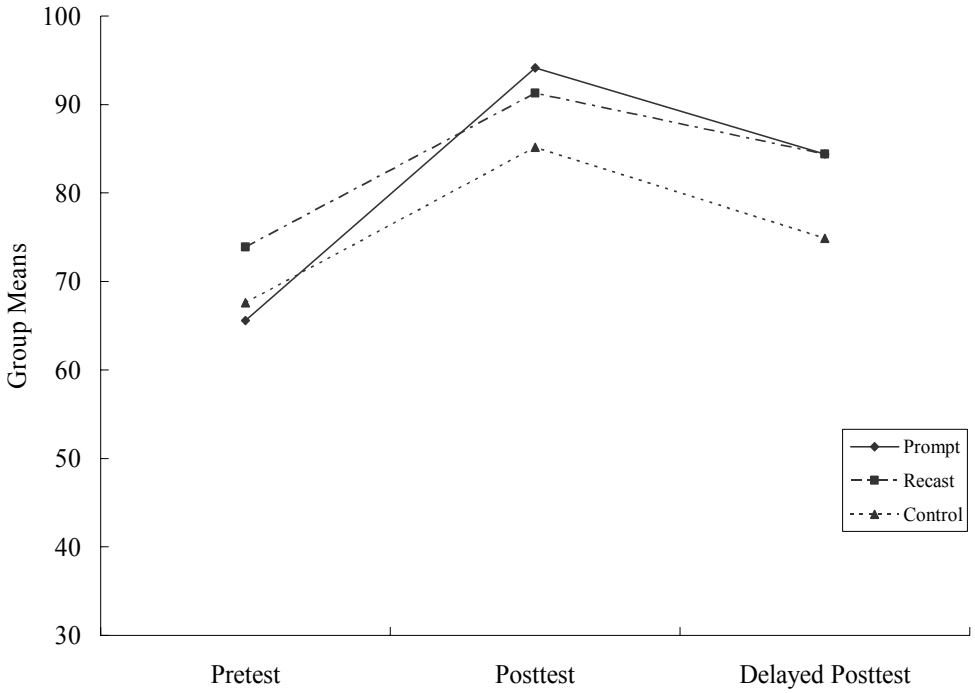


Figure 5.6. Group means on the use of regular past tense in the written test

5.2.3 Summary of Results from Quantitative Analysis

This section summarizes the results of the quantitative analysis with a view to answering the three research questions proposed at the end of Chapter 3. In order to answer these questions, it is necessary to revisit the research hypotheses. Each hypothesis is restated in turn and immediately followed by a summary of the corresponding results.

Hypothesis 1: Both the prompt group and the recast group will outperform the control group on both oral and written tests at the time of the post-test and the delayed post-test.

With respect to the overall benefit of feedback in the acquisition of past tense forms, it is difficult to come to a firm conclusion that the two treatment groups outperformed the control group. This is due to the fact that the results from the above quantitative analysis show that the two feedback groups did not perform unequivocally better in terms of the acquisition of past tense forms. Rather, all three groups improved to a great extent from the pre-test to the post-test and maintained this increase at the time of the delayed post-test in both the oral and the written test, as demonstrated by an overall highly significant time effect. The only indication of the superior effect of feedback was in the written test results, which revealed that the prompt group significantly outperformed the control group in the use of overall past tense forms as well as irregular past tense forms at the time of the post-test. The recast group did not distinguish itself from the control group, as reflected by multiple comparisons of post-test scores at each time point. However, in terms of the use of irregular past tense in the written test, post hoc comparisons did reveal an advantage in feedback, in that both the recast and the prompt group significantly improved from the pre-test to the post-test as well as from the pre-test to the delayed post-test, in contrast to the control group, which did not show any significant increase in scores. As an overall result, therefore, hypothesis 1 was partially confirmed.

Hypothesis 2: Prompts will have more beneficial effects than recasts in learning regular past tense forms (rule-based structure) measured by both the oral and the written post- and delayed post-tests.

In terms of the beneficial effect of prompts over recasts in learning regular past tense forms, the ANOVA results of the oral tests failed to show any advantage of prompts over recasts. However, Figure 5.3 graphically demonstrated that while the recast and the control groups maintained their increase in mean scores at the time of the delayed post-test, the prompt group continued to improve its mean score. In the written test, post hoc tests showed that the three groups significantly improved their scores from the pre-test to the post-test, yet only the prompt group maintained this increase at the time of the delayed post-test. This again confirmed that the prompt group had the advantage of maintaining the increase in scores more than the recast group. Therefore, the answer to the second question is affirmative, in the sense that prompts seem to have more beneficial long-term effects than recasts in the learning of regular past tense forms.

Hypothesis 3: Recasts, because they provide positive evidence, will better assist in learning irregular past tense forms (item-based structure) measured by both the oral and the written post- and delayed post-tests.

Finally, with regards to the acquisition of irregular past tense forms, the results of the present study contradict the prediction of the hypothesis. In the written test, the ANCOVA result showed that the only significant difference in the post-test scores was between the prompt and the control group. The prompt group demonstrated superiority in the acquisition of irregular past tense as represented by a larger increase in accuracy scores at the time of the post-test. Although the recast group also significantly improved from the pre-test to the post-test, as well as from the pre-test to the delayed post-test, this improvement failed to show any significant superiority over the prompt group. In the oral

test, although no significant differences were found between the three groups, group means revealed that the prompt group better maintained the gain in scores at the time of the delayed post-test than did the recast and the control groups. Contrary to the prediction proposed in the hypothesis, the results distinguished the prompt group from the recast and the control group in its overall larger gain in the written test at the time of post-testing.

5.3 Results of Analysis of Questionnaire Data

While the first main division of the chapter focused on the quantitative analysis of testing scores, this section presents the results from the analysis of the questionnaire data. First, the learners' answers to the two questions in the background information questionnaire (see Appendix B for the complete questionnaire) are examined. These questions relate to the students' strategies of learning spoken English and their perceptions of the important aspects in spoken English. Second, the learners' answers to question 1 as well as excerpts of students' short answers to questions 2 – 4 in the exit questionnaire (a complete questionnaire can be found in Appendix B) are analyzed and presented.

5.3.1 Results of Analysis from the Background Questionnaire

The background information questionnaire was administered prior to the experiment and the pre-tests. Altogether, a total of 99 students answered the questionnaire. In addition to ascertaining the students' background information, this questionnaire also contained two questions related to the students' strategies of learning spoken English and their perceptions of the important elements in speaking good English.

In response to the question (*In which way do you practice oral English?*) six choices were provided: (A) with foreigners; (B) with my foreign English teacher only; (C) with my Chinese English teacher; (D) with my classmates; (E) with strangers in the English corner[22]; and (F) others (Please specify). The students were informed that they could tick multiple answers to the questions. Eighty-four students chose (D) in their answer to this question, indicating that practicing oral English with fellow students was the most common way. The second most-frequent choice was (C), practicing oral English with their Chinese English teacher with 75 students choosing this answer. About half (48) of the students chose either (C), (D), or a combination of (C) and (D), which indicated that they rely mostly on their classmates and Chinese teachers to practice oral English. Twenty-two students selected answer (A), which is practicing oral English with foreigners. Thirteen students reported that they practiced oral English with their foreign English teacher only; 11 students, that they practiced oral English with strangers in the English corner; and 10 students, that they preferred other ways of practicing oral English, such as watching movies, listening to cassettes or memorizing English texts.

[22] "The English corner" is a special kind of activity to practice oral English. English learners of various proficiency levels and sometimes native speakers of English gather together at "the English corner" to communicate in English. The venues can be in or out of a university campus.

With respect to the second question (*Which aspects do you think are important in oral English?)* the following possible answers were provided: (A) pronunciation; (B) vocabulary; (C) grammar; (D) the use of idiomatic expressions; (E) clarity of meaning; (F) fluency. The students' response was a combination of (A) and (F), indicating that they considered pronunciation and fluency as important factors in speaking good English. Forty-two students chose "clarity of meaning" and 38 students chose "vocabulary" as important factors in spoken English, while 18 students selected "the use of idiomatic expressions" and only 12 students chose "grammar". It is interesting to note that grammar was the least popular selection among all the six choices provided to the students.

5.3.2 Results of Analysis from the Exit Questionnaire

The exit questionnaire was administered immediately following the first post-test. The purpose of this questionnaire was to ascertain the participants' focus of attention during the treatment sessions and their opinions of the communicative activities as well as their attitudes towards feedback. There were four questions in this questionnaire, including one multiple-choice question (Question 1) and three short-answer questions. The short-answer questions were constructed in an open-ended manner so as not to restrict the scope of students' answers, but the wording of the questions was succinct enough to elicit clear-cut answers. A total of 94 participants answered this exit questionnaire. The following analysis was conducted based on the information provided by these 94 sets of responses.

In question number one (*Now that you have completed the tasks and the tests, what do you think they were all about?*), four choices were provided to the students: (A) they were practicing and testing writing; (B) they were practicing and testing my grammar; (C) they were practicing and testing listening and speaking; and (D) they were practicing and testing my vocabulary. The students were told that they could choose multiple answers to this question. Out of the four choices, 83 participants chose answer (A); that is, they considered these tasks and tests to be about listening and speaking. Ten students chose writing and 10 students chose grammar as the focus of these activities and tests. Finally, 7 students selected vocabulary as the focus of these tests and activities.

Question number two was construed as another way of asking about the focus of these activities and tests; however, it opened up the scope of the answers by not actually referring to any specific aspects of these activities and tests. In answer to question number two (*Please write a short paragraph saying what you think you learned from this experience),* as expected, the students' answers varied to a large extent. The researcher reviewed and categorized these answers according to their common themes. Out of these highly diverse answers, seven categories were identified. Due to the fact that students' answers were in the form of a short paragraph, overlapping themes may have emerged from their answers. In that situation, each theme was counted as an individual occurrence. A single student's answer to one question may have therefore contained multiple themes. The most frequently mentioned themes were "improvement in skills of dictation" and "improvement of listening skills". Thirty-three participants suggested that they improved their dictation skills in certain ways and 32 participants responded that these activities improved their listening skills. Twenty-six participants indicated they also

enlarged their vocabulary and 26 participants reported that they improved their speaking skills. Seventeen participants reported that they learned some aspects of grammar. Out of these 17 participants, 6 of them suggested a focus on tense and two of them explicitly indicated "past tense" as the focus of their learning. Seventeen students also mentioned that they realized the limitation in their English skills from this experience. Twelve students reported that they learned how to do team work with their learning partners. Other responses identified a variety of themes including assistance in exam preparation, writing skills, provision of a relaxed environment and morals of stories, to mention just a few. Due to their low frequency of occurrence and idiosyncratic nature of these answers, these categories were lumped into one large category as "others". There were 27 instances of answers in this category.

Question number three was aimed at investigating the participants' views of the activities and tests employed in the present study. In response to the first part of question three (*Do you like the tasks you performed in class?*), 60 participants answered affirmatively, indicating that they liked the activities and tests in the present study; 24 participants said they didn't like the activities and tests; the remaining 10 participants indicated they neither liked nor disliked the activities (see Appendix B for a complete table of participants' answers). Out of the 60 participants who answered "yes" to this question, their reasons were explained as follows: 31 participants thought that these activities and tasks could improve their English skills; 31 participants indicated that they considered the tasks interesting and motivating; 13 participants believed that these activities could create a relaxed atmosphere for learning in the classroom; 2 participants explained that they liked the activities because they enabled them to be more flexible in

terms of using English skills; and 1 student suggested that the activities were useful because they involved every student in the classroom.

Participants who indicated that they did not like the activities did so due to the following reasons: 14 participants thought that the tasks were too difficult for them; 9 participants felt that the tasks took too much time to do yet had little effect in their learning; 9 participants thought that the tasks were not interesting; and 2 participants indicated that the tasks required only a small group of people to speak.

Question four asked participants about their attitude towards feedback and why they liked or disliked being corrected in class. In answer to the first part of this question (i.e., their attitude towards feedback), 75 participants (80%) indicated that they liked to be corrected while they were speaking in class, whereas 19 participants (20%) indicated that they did not. In explanation of their positive attitude towards feedback, 61 participants expressed the idea that feedback could help them to improve their English skills, while 10 participants indicated that others may see their problems better. Interestingly, despite their overall positive attitude towards feedback, 9 participants also expressed their concerns of feeling embarrassed while being corrected in front of fellow students, and 9 participants preferred to be corrected after their completion of a sentence rather than be interrupted before they completed a whole sentence in their speech. Among the participants who answered that they did not like to be corrected while they were speaking, 14 indicated that feedback may interrupt their thought or their speech; 4 participants expressed their fear of loss of face in front of their fellow students; and 4 participants thought that mistakes should not be corrected in oral English at all because they just cared about the communication of meaning.

To summarize results from questionnaire data, the background information questionnaire revealed that the participants in the present study mostly chose to communicate with their Chinese English teachers and their fellow students as a preferred way to practice spoken English. In addition, they believed that the two most important elements in speaking good English were pronunciation and fluency. Analysis of exit questionnaire data revealed that a majority of participants in the present study thought the activities and tests were practicing listening and speaking skills. The participants' attention was not drawn specifically to grammar or past tense. Instead, because of the communicative nature of the activities, they indicated that they learned other aspects such as speaking, vocabulary, skills in dictation, and listening. Participants also demonstrated an overall appreciation of the activities and expressed a generally positive attitude towards feedback.

Chapter Summary

This chapter presented the results of the present study from different perspectives, including the analysis of feedback and repair moves from classroom transcripts during the treatment sessions, the quantitative analysis of learners' performance over time, as well as the results from the questionnaire data analysis. The next chapter discusses the outcomes of the above analyses, interprets the research findings in light of previous empirical research, identifies the limitations and implications of the present study and, finally, outlines directions for future research.

CHAPTER 6：DISCUSSION AND CONCLUSION

The previous chapter presented the data analysis and results of the study and revisited the hypotheses raised in Chapter 3. This chapter focuses on the examination of the findings in relation to those obtained in previous studies and explains the results in this research context. The chapter also discusses the findings with respect to their theoretical and pedagogical implications, points out strengths and limitations of the present study, and, finally, outlines directions for future research on corrective feedback in SLA.

6.1 Research Findings and Interpretations

The first research question in the present study asked whether the groups that performed communicative activities while receiving feedback showed an overall superiority in learning regular and irregular past tense over the control group, which did not receive feedback.

Research findings of the present study indicate that all three groups significantly increased their scores over time. The effect of these form-focused communicative activities had a remarkable impact in drawing the learners' attention to the target form and，therefore，led to a large gain in accuracy scores among all three groups.

Form-focused instruction has been shown to be an effective method of integrating grammar instruction with communicative tasks in a variety of contexts

(Harley, 1998; Ishida, 2004; Muranoi, 2000; Lyster, 1994; Swain & Lapkin, 1998, 2002). The present study took place in a form-oriented EFL context with a strong focus on grammar and translation. An informal interview with the teachers who participated in the study revealed that the intensive reading class, which is the core class in the English major program throughout the four years of university study, relied heavily on the teachers' explanation of grammar points, translation exercises, as well as reading comprehension and vocabulary exercises. The teachers use a combination of Mandarin and English as the medium of instruction. In this type of form-oriented class, these communicative activities provided ample opportunities for students to interact with their teachers and fellow students in class, which highly motivated students to draw their attention to the content of these activities and focus on target forms. This finding supports the counterbalance hypothesis proposed by Lyster and Mori (2006), which states:

> Instructional activities and interactional feedback that act as a counterbalance to the predominant communicative orientation of a given classroom setting will be more facilitative of interlanguage restructuring than instructional activities and interactional feedback that are congruent with the predominant communicative orientations. (p. 294)

As suggested by this hypothesis, the shift of attentional focus as a result of a different kind of instructional intervention may lead to interlanguage restructuring. This shift of attentional focus may also strengthen the connections between changes in long-term memory and actual language production. In the present study where there was an overall form-focused orientation, these communicative activities with a focus on certain forms drew the learners' attention away from a decontextualized focus on forms, which seemed to result in interlanguage restructuring in a productive manner.

The lack of significant differences between the feedback groups and the control group at each point in time may be due to the design of the study. Although students in the control group did not receive any feedback during the activities, they also performed the activities with their teachers and had the opportunity to practice the target forms during the activities. Furthermore, for ethical concerns, in the dictogloss activities, students in the control group received the original version of the texts (with all the target forms enhanced in bold) and were asked to compare their own texts with the originals. In the question-and-answer activity, their teacher summarized some of the mistakes at the end of the activity and handed out the target form list with both the base forms and the simple past tense forms, whereas the feedback groups only received feedback treatment during this activity. These form-focused activities in the control group may account for the fact that the difference between the feedback group and the control group[23] was not found in all of the measures in the use of regular and irregular past tense forms.

Despite the fact that the control group was not in the real sense a control group that did not receive any treatment at all, the feedback groups did demonstrate superior performance over the control group in the following aspects. First, the prompt group outperformed the control group in the use of overall past tense forms, and the use of irregular past tense forms at the time of post-testing in the written production tasks. Second, post hoc comparisons of the results of written tests revealed that both the recast and the prompt group significantly improved from pre-test to post-test, as well as from pre-test to delayed post-test in the use of irregular past tense forms, whereas the control group did not show any significant increase in scores.

[23] As mentioned in Chapter 4, the control group in the study also performed the communicative activities because the purpose of the study was to examine the effect of feedback per se instead of the effect of feedback together with form-focused communicative activities.

In all, the answer to the first research question confirmed results from previous studies that showed the beneficial effect of feedback in comparison with control groups (Tomesello & Herron, 1988, 1989; Mackey, 2006; Russell & Spada, 2006). Moreover, the overall significant time effect supports the counterbalance hypothesis (Lyster & Mori, 2006), indicating that form-focused communicative activities are effective in drawing learners' attention to the target form, and in the case of this study, maintaining better control of the use of the target form in the Chinese EFL context.

6.1.1 Recasts and Prompts

The second and the third research questions addressed the relative efficacy of recasts and prompts in the acquisition of regular and irregular past tense forms. Taken together, findings of the present study reveal an overall beneficial effect of prompts over recasts in the acquisition of past tense forms. This result is congruent with findings of a number of previous quasi-experimental studies in a variety of communicative contexts, such as French immersion in Canada (Lyster, 2004b), Canadian ESL (Ammar & Spada, 2006), and New Zealand ESL (Ellis et al., 2006). These studies were unanimous in their evidence of the overall beneficial effect of prompts over recasts in the acquisition of the target features.

In terms of learning rule-based regular past tense forms, results showed that prompts have more beneficial long-term effects than recasts. The superiority of prompts over recasts in assisting the acquisition of rule-based regular past tense forms was demonstrated mostly on the delayed post-test where the prompt group maintained the

gain in the written test scores on the post-test better than both the recast and the control group. The oral test results did not show any significant differences among the three groups; however, the graph of the mean scores illustrated that the prompt group continued to improve at the time of the delayed post-test whereas the recast and the control group only maintained their increase at post-test.

With respect to the learning of item-based irregular past tense forms, the results clearly showed that, in the written test, the prompt group outperformed the control group at the time of post-testing. Although the recast group also significantly improved over time, this improvement failed to show any superiority in comparison with the prompt group or the control group. In the oral test, although no significant differences were found among the three groups, group means revealed that the prompt group maintained the gain in scores at the delayed post-test better than the recast and the control group. Contrary to the prediction that the recast group would outperform the control and prompt group in the use of irregular past tense forms, it was the prompt group that distinguished itself from the recast and the control group on the written test by an overall greater gain. This finding supports results from previous studies also showing a more beneficial effect of prompts over recasts. While these studies mostly targeted rule-based grammatical structures (Ammar & Spada, 2006; Ellis et al., 2006, 2007), the results of the present study extend the scope of the structures that may benefit from prompts more than recasts to include item-based structures as well.

Possible explanations of these results may be construed in terms of the saliency and explicitness of recasts in relation to features of the target grammatical structure and opportunities for modified output, as discussed in the section that follows.

6.1.2 Saliency and Explicitness

Findings of the present study did not show any superiority of the recast group in the acquisition of item-based irregular forms over the prompt group, thereby contradicting the third research hypothesis. These contrary findings are congruent with a series of quasi-experimental studies that showed the superiority of prompts in general in comparison with recasts in the acquisition of morphosyntactic structures. The findings of the present study also support the claim that prompting engages students at a deeper level of processing than do recasts (de Bot, 1996, 2000; Lyster, 2002, 2004b).

In order to examine the reason for the overall differences between the feedback conditions, audio-recordings and classroom transcripts were analyzed with the belief that they may provide further insights regarding the characteristics of recasts within this specific classroom context. A number of studies that have investigated characteristics of recasts have demonstrated that the implicitness of recasts varies according to context, which could explain discrepancies in the results between laboratory and classroom settings. Ellis et al. (2006) argue that recasts can only assist acquisition if learners are consciously aware that changes have been made to their original utterances. This may not always be possible in all circumstances. The ambiguity of recasts in communicative and immersion classrooms has been discussed in a number of studies (Lyster, 1998a, 2002). As argued in Chapter 2, in order for recasts to be perceived as negative evidence, and therefore to create the opportunity for learners to make cognitive comparisons between their interlanguage and the target language, recasts need to be salient enough in the

course of communication (cf. Ellis & Sheen, 2006; Sheen, 2007). Some characteristics that are associated with saliency of recasts are shortness in length, frequency, intensity (Ellis & Sheen, 2006), fewer changes, the targeting of a single structure, enhanced discourse features such as stress (Doughty & Varela, 1998; Han, 2002), and occurrence in dyadic interaction context (Han, 2002; Long et al., 1998; Mackey & Philp, 1998).

Classroom recordings of the present study revealed that not all of the recasts bore these characteristics. For example, recasts provided during the dictogloss tasks in the present study often involved multiple changes with no stress or specific emphasis on the error. In addition, these recasts appeared in a classroom context while students were performing communicative activities in which the communication of meaning was the main focus. During these activities, the teacher in the recast group recast students' erroneous utterances in a similar manner as when repeating students' correct utterances, as shown in examples 6.1 and 6.2. These examples illustrate how the teacher responded to the students when they made an error or when simply following a correct sentence. In example 6.1, the teacher repeated the student's correct sentence so that the whole class could hear the correct version of the sentence during the dictogloss activity. In example 6.2, however, the teacher recast the erroneous utterance (the correction was not on past tense, but on the use of preposition and article), and continued with the next sentence that the learner failed to supply. Due to the minor change the teacher made to the original sentence and the processing demand of the next sentence, the learners may not have noticed the correction made by the teacher. Instead, it is very likely that they may have thought that the teacher was simply repeating the student's previous sentence and thus focused only on the missing sentence.

Example 6.1

S: We turned out everything in the basket. (Correct utterance)

T: We turned out everything in the basket. (Repetition)

Example 6.2

S: We looked at the picture of tin. (Error)

T: We looked at the picture **on the tin**, we thought of the juice… (Recast)

Furthermore, the teacher provided recasts on a variety of learner errors in addition to past tense errors, as seen in example 6.3. In this example, the learner made multiple errors in his/her original sentence by missing a verb and failing to use the past tense for the two verbs. The teacher corrected the whole utterance by adding the missing verb, supplying the missing adverbial phrase of place "at the palace", and correcting the errors on the use of past tense. All these alterations occurred in one sentence and therefore may not have drawn the learners' attention exclusively to the target forms. These findings are in line with many studies which have demonstrated that teachers' intentions and learners' perceptions of recasts may not always coincide (Sheen, 2007; Mackey et al., 2000; Mackey, 2006). More particularly, some studies have pointed out that long recasts with multiple errors are especially difficult for learners to perceive (Loewen & Philp, 2006; Philp, 2003; Sheen, 2006).

Example 6.3

S: When Cinderella at the ball, the prince set eyes on Cinderella and walk to her, ask her to dance.

T: When Cinderella **entered the ballroom** at the palace, the prince set eyes on

Cinderella. Walking over to her, he bowed deeply and **asked** her to dance. (Recast)

All these examples constitute evidence that recasts in the present study may share some similar traits as those that appear in communicative and immersion classrooms, where "…recasts of ill-formed utterances and repetitions of well-formed utterances together appear to confirm or disconfirm the meaning of a learner's message, not its form" (Lyster, 2004b, p. 404). Learners may have been confused about the intention of these recasts or experienced trouble locating the error even in such a form-oriented context. In other words, they may not have been aware of whether the teachers were recasting on the accuracy of meaning or form, nor did they seem able to differentiate between a recast on erroneous utterances and a repetition of correct utterances.

There are studies that have indeed demonstrated the positive effect of recasts on the acquisition of past tense forms. These empirical studies all share something in common: the recasts were provided with some kind of emphasis (Doughty & Varela, 1998), were consistently focused on one target form (Han, 2002), or were used in a laboratory context during negotiated interactions (McDonough, 2007). As noted by Sheen (2007), the majority of studies that have provided evidence for the beneficial role of recasts in the acquisition of grammatical features were carried out in laboratory settings, where the learner received one-on-one treatment on the grammatical structure. The implicitness of recasts in those studies was therefore greatly reduced. Lyster and Izquierdo (2010) also found that recasts and prompts had a similar effect in the

acquisition of French grammatical gender by adult FSL learners in a laboratory context. Quasi-experimental studies, however, have either shown little effect of recasts, or less effect of recasts in comparison with prompts (e.g., Lyster, 2004b; Ellis et al., 2006). The results of the present study lend support to the claim that recasts convey a certain degree of ambiguity during communications with a focus on meaning (Lyster, 1998a). Furthermore, the findings of the present study indicate that even in a form-oriented EFL context, without enhanced saliency or extra effort in making recasts explicitly corrective, the effect of recasts is limited when compared to prompts.

In contrast, the pedagogical purpose of prompts is much more salient and thus the corrective nature of prompts is much easier for learners to perceive. By definition, prompts withhold correct forms (and other signs of approval) by clearly indicating that something is wrong in learner utterances, thus leading learners to modify their responses (Lyster, 2004b). In this way, learners not only clearly notice their problems, but also rely on their own resources to retrieve the target forms. By responding to learners with clarification requests or elicitation techniques following their erroneous utterances, teachers push learners to reanalyze their internalized forms as well as their underlying systems. This reanalyzing and restructuring process is crucial for learners to form new hypotheses regarding the target language and modify their output in a more accurate way. Furthermore, in some instances of the present study, the learners were provided with metalinguistic information regarding the well-formedness of their utterances (see example 6.4). As shown in this example, the ambiguity of the clarification request was eliminated by another prompt (metalinguistic information) which resulted in a successful learner repair. As argued by Ammar and Spada (2006), metalinguistic clues may help

learners to identify the nature and locus of the error. In addition, these clues can also help learners to self-repair their errors, especially errors for which they have metalinguistic knowledge, but over which they may lack control in oral production.

Example 6.4

S: We turn out the bag.

T: We what? (Clarification request)

S: We turn out the bag.

T: What tense should we use here? (Metalinguistic information)

S: turned out the bag.

Given the above analysis of the differences in the degree of saliency and explicitness of the two types of feedback, it is not difficult to understand that learners' performance in the prompt group showed an overall superiority in comparison with the control group. The recast group, however, did not demonstrate better achievement in terms of test scores in comparison with the control group at each time point, as had been expected. Since the recasts operationalized in the present study were a mixture of both implicit and explicit type, involving one or more changes, with or without emphasis, it is premature to conclude that recasts are in general less effective than prompts in second language learning. Rather, the evidence in the present study suggests that recasts provided during the communicative activities are not salient enough to play a facilitative role in comparison with prompts in the learning of past tense forms in this particular classroom context. More empirical studies need to be conducted that would compare fine-grained sub-categories of recasts in comparison with one type of prompt.

6.1.3 Opportunities for modified output

Prompts and recasts differ from each other not only in terms of saliency and degree of explicitness, but also in terms of the opportunity for modified output. Modified output has been an essential component in theories of language acquisition (Long, 1996; Swain, 1993). In the output hypothesis, Swain (1993) argues that learners need to be pushed to make use of their own resources and stretch their linguistic abilities to their best. By producing modified output, learners are able to achieve higher levels of accuracy and fluency in their subsequent speech. Modified output can also contribute to second language development as suggested by Levelt's speech production model (Levelt, 1989; Izumi, 2003; McDonough, 2005). As Izumi (2003) points out, when learners modify their output, they either generate new output or reprocess their original output, both of which trigger additional grammatical encoding.

Although the issue of whether uptake and repair (or modified output) is indicative of or necessary for subsequent learning is still under debate (Mackey & Philp, 1998; McDonough, 2007), there have been a number of empirical studies that have demonstrated the relationship between modified output and successful learning. In William's (2001) study, the relationship between uptake and subsequent L2 development was investigated by associating Language Related Episodes (LRE) with the tailor-made test scores after the treatment. It was found that, when repair occurred, language development took place. Another study that showed the effectiveness of uptake in predicting L2 development is Loewen's (2005) study, which investigated the

effectiveness of focus-on-form instruction on subsequent learning in ESL classrooms in New Zealand. Loewen found that "successful uptake" was significantly related to gains in post-test scores in vocabulary and grammar. McDonough (2005) investigated the impact of negative feedback and learners' responses on ESL question development in a Thai EFL context. Analysis of test data revealed that the only significant predictor of ESL question development was the production of modified output involving developmentally advanced question forms.

Results of the present study suggest that recasts were indeed less likely to elicit uptake and repair in comparison with prompts. This finding is congruent with a number of observational feedback studies in communicative or meaning-oriented contexts in which learners tended to respond more frequently following prompts than recasts (Lyster & Ranta, 1997; Panova & Lyster, 2002). As the analysis of classroom transcripts revealed, there was only one repair out of 24 feedback instances in the recast group, whereas in the prompt group, there were 26 repairs out of 28 feedback instances. This self-repair process is professed to allow learners to reconstruct their knowledge of the target structure, enhance the correct form in memory, as well as create a greater awareness of the rule or item in question by way of this re-analysis process (de Bot, 1996; Lyster, 2004b). Furthermore, as argued by Lyster and Mori (2006) and Ammar and Spada (2006), even if there were equal numbers of uptake following recasts and prompts, the overall beneficial effect of prompts over recasts would probably be found, since uptake following prompts always reflects certain levels of analysis and hypothesis reevaluation, whereas uptake following recasts might only be a sign of noticing or mere repetition of the target form (Ammar & Spada, 2006).

The ANOVA results of the present study clearly illustrated that prompts were overall more effective than recasts in subsequent language development. The superior effect for prompts over recasts was more obviously shown in the written test. It is possible that in addition to the different degrees of saliency and explicitness in the discourse function, this differential effect may have risen from the different opportunities for modified output following recasts and prompts.

6.1.4 Feedback and Grammatical Structures

The results of the present study highlight the different effects of prompts and recasts in the acquisition of regular and irregular past tense forms. Specifically, in the acquisition of irregular past tense forms, both feedback groups showed an advantage over the control group, as revealed by the post hoc analysis. Both feedback groups significantly improved their accuracy scores from pre-test to post-test, and from pre-test to delayed post-test, whereas the control group did not show any significant improvement. In addition, the prompt group distinguished itself from the recast group by outperforming the control group at the immediate post-test. In the acquisition of regular past tense forms, however, the pattern was different. All three groups significantly improved over time and this time effect overshadowed the effect of feedback, with all three groups significantly improving their scores from pre-test to post-test. The only difference lay in that the prompt group significantly improved from pre-test to delayed post-test, whereas the other two groups did not. Possible explanations of these results may include the inherent differences of the target grammatical structures as well as the learners' prior knowledge

of the target structures.

Previous feedback studies have revealed that the provision of different types of feedback may be associated with different types of grammatical errors. Furthermore, the nature of grammatical errors may determine the effectiveness of different types of feedback. For example, Mackey, Gass, and McDonough (2000) found that learners were able to perceive feedback relatively accurately following lexical, semantic, and phonological errors, but not necessarily to perceive feedback following morphosyntactic errors. In their study, morphosyntactic errors were usually recast and therefore it was not clear whether it was the ambiguity exhibited in recasts that hampered the perception of feedback as corrective or the morphosyntactic errors that were difficult for learners to perceive.

The target structures in the present study are regular and irregular past tense forms. By nature, regular and irregular past tense forms carry varying levels of saliency during the course of communication. Regular past tense has always been associated with low saliency, low communicative value and high regularity (Ellis, 2005; DeKeyser, 1998), whereas irregular past tense has been associated with high saliency and high frequency, but low regularity (Salaberry, 2000).

In the present study, both feedback groups showed significant gains from pre-test to post-test and maintained their gains at the delayed post-test in the acquisition of irregular past tense forms. In addition, the difference between the prompt group and control group was significant at the time of post-testing. The positive effect of recasts on the acquisition of irregular past tense forms suggests that the high saliency of the target form and obvious change made to the non-target forms rendered recasts easier for

learners to perceive as negative evidence during the course of communication. In contrast, recasts provided on the regular past tense forms may have only involved the addition of the morpheme *–ed*, which lacked saliency in comparison with changes made to the entire verb, or changes that involved vowel alternations in the case of irregular past tense forms. This may explain why the recast and the control group behaved similarly in the written test on the use of regular past tense forms. As Salaberry (2000) has argued, "…In essence, the prediction is that the more frequent and irregular the verb the more likely it will appear first in the development of past marking of adult instructed L2 learners" (p. 138).

In Mackey's (2006) study that investigated the relationship between feedback and noticing, out of the three different grammatical structures examined, question formation was the most noticeable structure, whereas regular past tense was the least noticeable. Mackey explained this result in terms of the saliency of the grammatical structures. She argued that since question formation involves syntactic movement as well as morphological agreement, it may be more noticeable than the past tense morpheme. Question forms are also more salient because of their higher frequency in classroom discourse and high communicative value. In the present study, feedback on regular past tense morphemes that lack saliency and communicative value may not have been as noticeable as irregular past tense forms. The fact that Chinese Mandarin does not mark past tense with morphological change (Cai, 2007) may also account for the lack of noticing of both the irregular and the regular past tense morphemes, since expectancy is one of the factors that may influence the noticing of grammatical features (Schmidt, 2000). The addition of the sometimes voiceless morpheme *–ed* makes regular past tense forms even less noticeable in the course of communication.

In addition, the superior performance of the prompt group over the recast group on the acquisition of the irregular past tense forms may be due to the fact that, by being pushed to provide modified output, learners had to retrieve the target forms themselves from their long-term memory. This retrieval process may have reinforced the connection of the item-based target form (i.e., irregular past tense) in memory, whereas in the case of the recast group, the learners did not engage in the same level of retrieval process as the prompt group learners (Lyster, 2002, 2004b). As a result, the prompt group showed better performance than the recast group in the use irregular past tense forms and, though less obvious, in the use of regular past tense forms[24] in the written test. This finding supports the claim that providing learners with only positive evidence, even in the learning of item-based target forms, is not as effective as engaging them in deeper levels of analysis and restructuring (de Bot, 1996; Lyster & Mori, 2006).

In his study that measured learning difficulty of a number of grammatical structures, Ellis (2005) found that regular past tense is among those features for which rules are applicable, easily instructed explicitly, yet difficult to be learned as implicit knowledge. The results of the written test of the present study, which arguably involved the use of explicit knowledge, support Ellis's claim by showing an overall higher accuracy rate for the use of regular past tense form than that of irregular past tense forms. The lack of significant group differences may lie in the fact that all the groups showed relatively higher pre-test scores (around 70%). As argued by Ellis (2007), the well-established explicit knowledge of past-tense *–ed* may imply a ceiling effect for prompts.

Another interpretation of the findings is that the form-focused communicative

[24] The means of the different groups reveal that the prompt group went from the lowest score at the pre-test to the highest at post-test, although this failed to result in a statistically significant difference.

activities triggered the learners' awareness of using the grammatical structure of past tense in general. This may have helped them in achieving a higher score on the use of regular forms irrespective of feedback conditions, since the rule of forming regular past tense is relatively easier to apply. However, with respect to the highly unpredictable, item-based irregular forms, even if learners were aware that they should use past tense forms on the irregular verbs, they did not have the resources in their repertoire to do so. For those item-based forms, learners may need special assistance, either by being pushed to retrieve the forms from their long-term memory, or by being provided with the target forms. The results suggest that pushing learners to retrieve the item-based forms led to better acquisition of the irregular past tense than did recasts.

A higher accuracy score on the use of regular past tense marking than irregular past tense marking was also found in Cai's (2007) analysis of Chinese EFL learners' interlanguage past tense marking. She explains that these Chinese foreign language learners were taught grammatical rules first, so they learned to differentiate irregular and regular verbs and how to mark each word individually before they generalized certain rules. In other words, the acquisition of regular and irregular past tense forms for Chinese EFL learners in classroom settings, similar with first language speakers, is through the dual-mechanism process proposed by Pinker and his colleagues (Pinker, 1999; Pinker & Prince, 1994), one involving the acquisition of a regular morphological rule, the other through associative learning of the irregular forms. As a result, learners in Cai's (2007) study and the present study may have had fair knowledge of the rule applying *–ed* to the regular verbs; however, it may have taken a longer time for them to memorize all the irregular past tense forms, for which consistent feedback and practice could have been

useful in providing assistance.

6.1.5 The Development of Implicit and Explicit knowledge

The oral test in the present study employed a testing measure to tap into the learners' implicit knowledge (Ellis, 2005; Ellis et al., 2006). Contrary to the written test results, the results of the oral test failed to show any statistically significant differences between the three groups. However, group means and plots of three groups' performance in the oral test over time did reveal that the prompt group seemed to maintain the gain in scores better than the control and the recast groups at the delayed post-test. On the use of regular past tense, the prompt group, in fact, continued to increase its scores at the delayed post test.

These findings can be interpreted according to the following aspects. First, the test used in the study requires complex on-line production at a discourse level instead of imitation or sentence level production (Brown, 2004). The oral narrative employed in the present study requires learners to draw on their memory of a text that they read for only three minutes and then retell the story based on word cues without any further planning time. This is a very challenging task for intermediate-level learners. The on-line processing demand may have deprived the learners of their reliance on explicit knowledge of the past tense. Second, drawing on generally different results between written and oral production tasks, Lyster (2004b) claims that it is possible that prompting affects online oral production skills to a lesser degree in comparison with written production. This may be due to the fact that prompts seem to increase students'

metalinguistic awareness and their ability to draw on declarative knowledge on tasks in which they have sufficient time to control their production. In the same vein, Ellis (2005) argues that deeply embedded knowledge (or implicit knowledge) requires automatic processing, whereas explicit knowledge requires controlled processing. The development of automatic or implicit knowledge may take a longer time than the development of explicit knowledge. Third, Ellis also argues that the use of the two types of knowledge depends on the tasks that learners perform. Yuan and Ellis (2003) illustrated that given plenty of time to plan production on-line, learners' speech becomes more accurate because they can access explicit knowledge. In contrast, when learners are required to perform the same task under pressure, their speech becomes less accurate. This may explain why the three groups performed similarly in the oral tests. Furthermore, Skehan (1998) claims that there is always a trade-off effect between fluency, accuracy, and complexity. It is possible in the present study, then, that high task demand, no planning time, and lack of conscious control over the target forms may have led to the three groups' similar gains in the accuracy score in the oral test.

Another factor that may explain the lack of group differences in the oral test is the short duration of the feedback treatment. The total treatment of the four communicative activities was 2 hours spread over 2 weeks. In previous studies, feedback treatment has ranged from 30 minutes (Loewen & Nabei, 2007) to four or five weeks (Ammar & Spada, 2006; Lyster, 2004b). The duration of the treatment in the present study may not have been enough to truly illustrate the effect of feedback, especially in the learners' oral production. As predicted by the ACT theory (Anderson, 1985; de Bot, 1996; DeKeyser, 1998), all kinds of knowledge need to undergo a transition from

declarative knowledge to procedural knowledge. The transition from the highly controlled declarative knowledge to less controlled, automatic procedural knowledge needs repeated practice as well as feedback (Lyster, 2002; Lyster & Izquierdo, 2010). The results of the present study have begun to show the advantage of prompts at delayed post-test over the recast and the control group in the oral test. This finding is in line with previous findings by Ellis et al. (2006) who also found the effect of prompts only in the delayed post-test. It is possible that, with a longer treatment, the effect of feedback would also be seen in the oral test. This hypothesis, however, would need to be empirically investigated in future research.

Taken together, the mixed results of the oral and written test show that the two measures tap into different types of knowledge. Furthermore, these results demonstrate that participants in the present study may have well-developed metalinguistic knowledge, as revealed by their high accuracy scores in the written test, yet lack control over the target form in the oral production. As Sharwood Smith (1986) argues, "some rule or principle may be acquired (in the competence sense) but suffer a long delay before full control is established" (p. 12).

6.2 Contributions and Implications

Along with previous studies on the differential effect of varying types of feedback, the present study supports the claim that prompts are more effective than recasts in the learning of rule-based grammatical structures (Ammar & Spada, 2006; Ellis et al., 2006; Ellis, 2007; Lyster, 2004b), and extends this claim to the acquisition of item-

based grammatical structures. These results also confirm previous findings regarding the positive overall effect of feedback (Tomosello & Harron, 1988, 1989; Mackey, 2006; Russel & Spada, 2006) and reinforce the importance of negative evidence (Gass, 2003) and noticing (Schmidt, 2001) in second language learning, especially for adult learners.

The present study also contributes to second language acquisition theory in the following ways. The results support Skehan's (1998) dual-mode hypothesis in that the acquisition of regular past tense and irregular past tense took place along different patterns. Significant differences between the groups were seen more in the acquisition of irregular past tense forms than in the acquisition of regular past tense forms. It is possible that the irregular past tense forms (item-based structures) are more amenable to feedback treatment than the regular past tense within the present context. Another interpretation could be that focus-on-form communicative activities triggered learners' awareness of the rule for past tense marking of the regular past tense; therefore, the effect of feedback could not be teased apart from the effect of the activities.

The different acquisition patterns also suggest that, for these highly-motivated EFL learners who have experienced extensive form-focused instruction, the learning of rule-based forms and item-based forms requires different kinds of treatment and practice. On the one hand, for the acquisition of partially acquired rule-based grammatical structures, learners need some kind of focus-on-form activities that allow them to actually use the language form in a communicative context. These activities would trigger their awareness of the application of the rules and allow them to gain greater control of the target forms in similar contexts. On the other hand, in the acquisition of partially acquired item-based grammatical structures, learners need both focus-on-form

activities and teachers' feedback to provide them with opportunity to use the language form in real communicative situations and, at the same time, to notice the discrepancy between their erroneous form and the target form. Prompts would assist them to reevaluate their hypothesis on the forms which may not be in a stable state in their short-term memory. Prompts also push learners to retrieve the target forms themselves from their long-term memory and repair their own erroneous forms, whereas recasts provide learners with the target forms without necessarily engaging them in the restructuring of their interlanguage.

The second important contribution made by the present study is that it extends the scope of feedback studies from communicative and immersion classrooms to the EFL context. While there have been a number of observational studies on feedback and uptake in foreign language contexts (Havranek, 2002; Lochtman, 2002; Ohta, 2000; Sheen, 2004; Tsang, 2004), there have not been many quasi-experimental studies in the EFL context that empirically compare the effect of different types of feedback especially with relatively large sample sizes. The present study took place in a typical form-focused EFL context in China with participants who were uniformly Chinese university students and shared similar learning experience and foreign language exposure. These factors make the intact groups in the present study comparable while maintaining the ecological validity of the study. As argued by Ellis et al. (2006), the findings of laboratory studies are not necessarily comparable to those obtained in classroom contexts. DeKeyser (2001) also claims that the findings in the laboratory studies need to be validated in classroom settings. In addition, the treatment was conducted by Chinese English teachers instead of native speakers of English, which may have considerable pedagogical implications for

teachers and researchers in a foreign language context. Researchers and teachers could adopt the research methods employed in the present study and conduct classroom research in other similar settings to validate the findings. In addition, teachers in the Chinese context could adapt the activities and tests developed for the present study to meet their pedagogical purposes and apply them in their teaching practices.

Another contribution of the present study is related to the design and data collection procedures. Unlike some quasi-experimental studies that compared the effect of different feedback groups with a control group that only participated in the tests (Ammar & Spada, 2006; Ellis et al., 2006; Ellis, 2007; Sheen, 2007), the control group employed in the present study performed all the activities as did the feedback groups, yet differed from the feedback groups only in the feedback treatment (also see Lyster, 2004b). This can be considered a strength of the present study because it allows the researcher to tease apart the effect of communicative activities from the effect of feedback. Although this design may not be as likely as other studies to reveal a significant difference between the feedback groups and the control group, the results can nonetheless profess to reveal more convincing evidence of the isolated effect of feedback.

One unique feature of the design of the present study is the inclusion of audio-recordings of the actual treatment sessions. Audio-recording allows the researcher to directly pinpoint the amount of error, feedback and uptake that actually occurs during the communicative activities across groups, so as to establish baseline data and guarantee that the amount of treatment between the two feedback groups is comparable. It also allows the researcher to oversee the implementation of different treatment conditions to ensure that the proper instructional procedure is being followed. The third advantage of

the inclusion of audio-recordings of the treatment sessions is that it enables the researcher to analyze the specific characteristics of the feedback provided in order to explain the differential effect of prompts and recasts in a more refined manner.

Finally, the inclusion of two test measurements of both oral and written modality also increases the reliability of the test results by tapping into learners' explicit and implicit knowledge (Ellis, 2005). Previous form-focused instruction and feedback studies that involved the use of oral and written tests have shown that the two types of tests yield different results (Lyster, 2004b). Lyster found in his study that, while the written test showed a clear effect of prompts, oral production did not distinguish between feedback conditions to the same extent.

6.3 Limitations and Future Research

The present study has certain limitations that need to be improved in future research. The first is the limited number of intact classes within each group and the possible influence of the teacher. The present study had only one class in each group and one teacher for each treatment condition. Although the treatment conditions were assigned randomly to the intact classes for each group, the limited number of classes could not eliminate the possible effect of the teacher. This challenge occurs in the majority, if not all, of educational studies. Ideally, the inclusion of two or more classes within the same school or across different school boards would increase the reliability of the study. However, there were only four classes at the same level in the same school in this study, which excluded this possibility. In future studies, if possible, increasing the

number of classes with comparable participants and teachers would provide more reliable and robust findings.

The second limitation is the short duration of the treatment. As mentioned in the previous section, the relative effect of prompts and recasts might have been demonstrated more clearly had the treatment sessions been longer. However, due to the constraints of the availability of the participants and their teachers, the treatment sessions could not last any longer than two hours. The researcher conferred with the teachers and decided that the time of the treatment sessions should be commensurate with the curriculum with respect to the focus of the target features, so as to find a natural period for the inclusion of the communicative activities and feedback treatment. Future studies that compare the effect of prompts and recasts with longitudinal designs and with a wider range of grammatical structures may add current knowledge of the effect of feedback in relation to grammatical structures to a greater depth.

Another limitation of the study is that the tests employed to measure the learner's development of target forms were not counterbalanced in the present study. All participants in the three classes underwent the same sequence in the oral and written tests, although there were minor variations of the test versions. This may have rendered the results comparable, yet the test-retest effect could not be eliminated. In addition to the non-counterbalance issue, the written test only included a limited number of target forms. Future studies may include counterbalanced tests with more than one test measure in each modality (oral and written), each including a substantial amount of target structures in order to triangulate the results.

The present study was designed according to the assumption that past tense

morphology involves the binary classification of rule-based regular past tense and item-based irregular past tense (Pinker & Prince, 1994; Pinker, 1999). Previous studies have revealed that past tense morphology has other inherent properties such as phonological variations (Bayley, 1991, 1994; Wolfram, 1985) which may affect the saliency of past tense marking. Other researchers have hypothesized that learners in naturalistic settings acquire past tense through a rather fixed order depending on the semantic features of different verbs, a hypothesis most commonly known as the tense-and-aspect hypothesis (Andersen & Shirai, 1996; Shirai, 1991). First language transfer may also play a significant role in the acquisition of past tense (Cai, 2007). Cai argues that interlanguage variation reveals systematic development from learners' first attempt to use the target language to more advanced levels as they approach further towards the target language. As a result, he considers interlanguage as an important indication of different stages in the language acquisition process. A systematic analysis of Chinese EFL learners' interlanguage on the use of past tense by Cai (2007) reveals that interlanguage variation is not only related to a number of interlinguistic and intralinguistic factors such as verb saliency, tense and aspect, first language influence and narrative structure, but also that these factors are interrelated to influence interlanguage variation. Future studies that examine these different factors in relation to feedback in affecting the acquisition of past tense would indeed shed light on the acquisition of past tense morphemes.

As commonly acknowledged in all studies, due to the limitations of the present study, one needs to take great caution in generalizing the results of the present study to other settings, to participants of different characteristics, or to the acquisition of other grammatical features. In this regard, future research should address the issue of

differential effects of recasts and prompts in the acquisition of a wide variety of language structures, with wider populations in different contexts. The present study assumes that participants were at the intermediate level, but did not further investigate the effect of feedback in relation to individual differences such as their proficiency level (Ammar & Spada, 2006), language analytic ability and aptitude (Sheen, 2007), level of motivation to learn English (Dörnyei, 2001), or cognitive factors such as working memory (Trofimovich, Ammar, & Gatbonton, 2007). Nor were there any retrospective data (Egi, 2007b; Mackey, 2006) that could directly illustrate the relationship between learners' perception of different types of feedback and feedback efficacy in learning the target structure. These are all interesting areas for further investigation.

6.4 Conclusion

The present study was motivated by the theoretical debate as well as practical concern over which type of feedback is more effective in second language acquisition. Drawing on a substantial amount of research that has established the positive effect of prompts over recasts on rule-based grammatical structures (Ammar & Spada, 2006; Ellis et al., 2006; Ellis, 2007; Lyster, 2004b), the present study set out to investigate the effects of different kinds of feedback on different kinds of grammatical structures in an EFL context (Ellis, 2007).

The results of the study support previous claims that prompts have an overall advantage in assisting the acquisition of rule-based grammatical structures. The results also provide empirical evidence that prompts are superior in the acquisition of item-based

structures as well. In addition, the findings confirm that the extent to which learners are able to benefit from feedback depends, in part, on the characteristics of the target structure (Egi, 2007b; Ellis, 2007) and indirectly suggest that the inherent nature of recasts may mediate the effectiveness of recasts in promoting acquisition.

Although the results of this study generally demonstrate the advantage that prompts hold over recasts in the acquisition of the target structure, it is necessary to take caution in interpreting and generalize the findings. As noted by Ellis (2007), it cannot be concluded that prompts would be more effective than recasts in the acquisition of all grammatical structures. It is necessary to add here that the same results may not be found in other contexts with other participants. Great care needs to be taken in future studies in the operationalization of the feedback treatment, the design and implementation of the procedures, and the selection of target structures. More research is needed to validate the results of the present study and to extend the scope of the grammatical structures that are amenable to different feedback treatment.

REFERENCES

Ammar, A. (2003). *Corrective feedback and L2 learning: Elicitation and recasts.* Unpublished doctoral dissertation, McGill University, Montreal.

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

Andersen, 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 of second language acquisition* (pp. 527-570). San Diego: Academic Press.

Anderson, J. (1980). *Cognitive psychology and its implications*. San Francisco: Freeman.

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

Anderson, J. (1985). *Cognitive psychology and its implications*. (2nd ed.) New York: W. H. Freeman and Co.

Anderson, J., Fincham, J. & Douglass, S. (1997). The role of examples and rules in the acquisition of a cognitive skill. *Journal of Experimental Psychology: Learning, Memory and Cognition, 23*, 932-945.

Bax, S. (2003). The end of CLT: a context approach to language teaching. *ELT Journal, 57*, 278-287.

Bardovi-Harlig, K. (1998). Narrative structure and lexical aspect: Conspiring factors in second language acquisition of tense-aspect morphology. *Studies in Second Language Acquisition, 20*, 471-508.

Bardovi-Harlig, K. (1999). From morpheme studies to temporal semantics: Tense-aspect research in SLA. *Studies in Second Language Acquisition, 21*, 341-382.

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

Bayley, R. (1991). *Variation theory and second language learning: Linguistic and social constraints on interlanguage tense marking*. Unpublished Doctoral Dissertation. Stanford University, Stanford, CA.

Bayley, R. (1994). Interlanguage variation and the quantitative paradigm: Past tense marking in Chinese English. In S. Gass, E. Cohen & E. Tarone (Eds.), *Research Methodology in Second Language Acquisition* (pp. 157-181). Hillsdale, NJ: Erlbaum.

Bolinger, D. (1975). Meaning and memory. *Forum Linguisticum, 1*, 2-14.

Brown, H. (2004). *Language assessment: Principles and classroom practices.* White Plains, NY: Pearson Education.

Cai, J. (2007). *The effects of multiple linguistic factors on the simple past use in English interlanguage.* Beijing: Foreign Language Teaching and Research Press.

Carr, T. H., & Curran, T. (1994). Cognitive factors in learning about structured sequence: Applications to syntax. *Studies in Second Language Acquisition, 16*, 205-230.

Carroll, S. (2001). *Input and evidence: The raw material of second language acquisition*. Philadelphia: John Benjamins.

Carroll, S. & Swain, M. **(**1993). Explicit and implicit negative feedback: An empirical study of the learning of linguistic generalizations. *Studies in Second Language Acquisition, 15*, 357-386.

Chaudron, C. (1977). A descriptive model of discourse in the corrective treatment of learners' errors. *Language Learning, 27*, 29-46.

Chaudron, C. (1988). *Second language classrooms.* Cambridge: Cambridge University Press.

de Bot, K. (1996). The psycholinguistics of the output hypothesis. *Language Learning, 46*, 529-555.

DeKeyser, R. (1993). The effect of error-correction on L2 grammar knowledge and oral proficiency. *Modern Language Journal, 77*, 501-514.

DeKeyser, R. (1995). Learning second language grammar rules: An experiment with a miniature linguistic system. *Studies in Second Language Acquisition, 17*, 379-410.

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: Cambridge University Press.

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

Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 129-154). New York: Cambridge University Press.

Dörnyei, Z. (2001). Teaching and researching motivation. Harlow, UK: Longman.

Egi, T. (2007a). Interpreting recasts as linguistic evidence: The roles of linguistic target, length, and degree of change. *Studies in Second Language Acquisition*, *29*, 511-537.

Egi, T. (2007b). Recasts, learners' interpretations, and L2 development. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 249-267). Oxford: Oxford University Press.

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

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

Ellis, R. (2005). Measuring implicit and explicit knowledge of second language: a psychometric study. *Studies in Second Language Acquisition, 27*, 141-172.

Ellis, R. (2006). Modelling learning difficulty and second language proficiency: The differential contribution of implicit and explicit knowledge. *Applied Linguistics, 27*, 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: A collection of empirical studies* (pp. 339-360). Oxford: Oxford University Press.

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meaning. *Studies in Second Language Acquisition, 21*, 285-301.

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Learner uptake in communicative ESL lessons. *Language Learning, 51*, 281-318.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition, 28*, 339-368.

Ellis, R. & Sheen, Y. (2006). Re-examining the role of recasts in L2 acquisition. *Studies in Second language Acquisition, 28*, 575-600.

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics*, 27, 464–491.

Fazio, L., & Lyster, R. (1998). Immersion and submersion classrooms: A comparison of instructional practices in language arts. *Journal of Multilingual and Multicultural Development, 19*, 303-317

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

Gass, S. M. (2003). Input and interaction. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 224-255). Malden, MA: Blackwell Publishing.

Gass, S., Mackey, A., & Pica, T. (1998). The Role of Input and Interaction in Second Language Acquisition: Introduction to the Special Issue. *The Modern Language Journal, 82*, 299-305.

Guthrie, E. (1984). Six cases in classroom communication: A study of teacher discourse in the foreign language classroom. In J. Lantolf & A. Labarca (Eds.), *Research in Second Language Learning: Focus on the Classroom* (pp. 173-194). Norwood: Ablex.

Han, Z. (2002). A study of the impact of recasts on tense consistency in L2 output. *TESOL Quarterly, 36*, 542-572.

Harley, B. (1998). The role of form-focused tasks in promoting child L2 acquisition. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 156-174). Cambridge: Cambridge University Press.

Harley, B., & Swain, M. (1984). The interlanguage of immersion students and its implications for second language teaching. In A. Davies, C. Criper, & A. Howatt (Eds.), *Interlanguage* (pp. 291-311). Edinburgh: Edinburgh University Press.

Hatch, E. (1978). Discourse analysis and second language acquisition. In E. Hatch (Ed.), *Second language acquisition: A book of readings* (pp. 401-475). Rowley, MA: Newbury House.

Havranek, G. (2002). When is corrective feedback most likely to succeed? *International Journal of Educational Research*, *37*, 255-270.

Havranek, G., & Cesnik, H. (2001). Factors affecting the success of corrective feedback. *EUROSLA Yearbook, 1*, 99-122.

Hu, G. (2003). English language teaching in China: Regional differences and contributing factors. *Journal of Multilingual and Multicultural Development, 24*, 290-318.

Hu, G. (2005). 'CLT is best for China' — an untenable absolutist claim. *ELT Journal, 59*, 65-68.

Hulstijn, J. (1990). A comparison between the information processing and analysis/control approaches to language learning. *Applied Linguistics, 11*, 30-45.

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.

Inagaki, S., & Long, M. (1999). Implicit negative feedback. In K. Kanno (Ed.), *The acquisition of Japanese as a second language* (pp. 9-30). Amsterdam: Benjamins.

Ishida, M. (2004). Effects of recasts on the acquisition of the aspectual form *–te i-(ru)* by learners of Japanese as a Foreign Language. *Language Learning, 54*, 311-394.

Izumi, S. & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly, 34*, 239-278.

Izumi, S. & Bigelow, M. (2001). Methodological and theoretical issues in testing the effects of focus on form. *TESOL Quarterly, 35,* 181-189.

Izumi, S. (2003). Comprehension and production processes in second language learning: in search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics, 24,* 168-196.

Kim, J., & Han, Z. (2007). Recasts in communicative EFL classes: do teacher intent and learner interpretation overlap? In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 269-297). Oxford: Oxford University Press.

Kowal, M. & Swain, M. (1994). Using collaborative language production tasks to promote students' language awareness. *Language Awareness, 3*, 73-93.

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

Krashen, S. (1985). The input hypothesis: Issues and implications. New York: Longman.

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

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.

Liao, X. (2004). The need for Communicative Language Teaching in china. *ELT Journal, 58*, 270-273.

Lightbown, P. (1998). The importance of timing in focus on form. In: C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 177-196). Cambridge: Cambridge University Press.

Lightbown, P., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition, 12*, 429-448.

Lightbown, P., & Spada, N. (1994). An innovative program for primary ESL in Quebec. *TESOL Quarterly, 28*, 563-573.

Lochtman, K. (2002). Oral corrective feedback in the foreign language classroom: how it affects interaction in analytic foreign language teaching. *International Journal of Educational Research, 37*, 271-283.

Loewen, S. (2002). The occurrence and effectiveness of incidental focus on form in meaing-focused ESL lessons. Unpublished Doctoral Thesis. The University of Auckland, Auckland, New Zealand.

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition, 27*, 361-286.

Loewen, S., & Nabei, T. (2007). Measuring the effects of oral corrective feedback on L2 knowledge. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 361-377). Oxford: Oxford University Press.

Logan, G. (1988). Towards an instance theory of automatisation. *Psychological Review, 95*, 492-527.

Long, M. H. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics, 4*, 126-141.

Long, M. H. (1985). Input and second language acquisition theory. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 377-393). Rowley, MA: Newbury House.

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), New York, CA: Academic Press.

Long, M. (2007). *Problems in SLA*. Mahwah, NJ: Erlbaum.

Long, M. H., Inagaki, S., & Ortega, L. (1998). The role of implicit negative feedback in SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal, 82*, 357-371.

Long, M., & Robinson, P. (1998). Focus on Form: Theory, research, and practice. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15-41). New York: Cambridge University Press.

Lyster, R. (1998a). Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition, 20*, 51-81.

Lyster, R. (1998b). Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, *48*, 183-218.

Lyster, R. (2002). Negotiation in immersion teacher-student interaction. *International Journal of Educational Research, 37*, 237-253.

Lyster, R. (2004a). Research on form-focused instruction in immersion classrooms: implications for theory and practice. *French Language Studies, 14*, 321-341.

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

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition, 19*, 37-66.

Lyster, R. & Mori, H. (2006). Interactional feedback and instructional counterbalance. *Studies in Second Language Acquisition, 28,* 269-300.

Lyster, R., & Izquierdo, J. (2010). Prompts versus recasts in dyadic interaction. *Language Learning*, *59*, XXX-XXX.

Mackey, A. (2006). Feedback, noticing, and second language development: an empirical study of L2 classroom interaction. *Applied Linguistics, 27*, 405-430.

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

Mackey, A., Gass, S., & McDonough, K. (2000). How do learners perceive interactional feedback? *Studies in Second Language Acquisition, 22*, 471-497.

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 (Eds.), *Individual differences and instructed language learning* (pp. 181-209). Philadelphia: John Benjamins.

Mackey, A. & Gass, S. (2005). *Second language research: Methodology and design*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mackey, A., & Goo, J. (2007). Interaction in SLA: a research synthesis and meta-analysis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407-452). Oxford: Oxford University Press.

McDonough, K. (2005). Identifying the impact of negative feedback and learners' responses on ESL question development. *Studies in Second Language Acquisition, 27*, 79-103.

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: A collection of empirical studies* (pp. 323-338). Oxford: Oxford University Press.

Mito, K. (1993). The effects of modeling and recasting on the acquisition of L2 grammar rules. Unpublished manuscript, University of Hawai‘i at Mānoa.

Muranoi, H. (2000). Focus on form through interaction enhancement: Integrating formal instruction into a communicative task in EFL classrooms. *Language Learning, 50*, 617-673.

Nelson, K. (1981). Toward a Rare Event Cognitive Comparison Theory of Syntax Acquisition. In P. Dale & D. Ingram (Eds.), *Child Language: An International Perspective* (pp. 375-402). Baltimore: University Park Press.

Nelson, K. (1987). Some observations from the perspective of the rare event cognitive comparison theory of language acquisition. In K. Nelson & A. Kleeck (Eds.), *Children's language* (Vol. 6, pp. 289-331). Hillsdale, NJ: Lawrence Erlbaum.

Nicholas, H., Lightbown, P., & Spada, N. (2001). Recasts as feedback to language learners. *Language Learning*, *51*, 719-758.

Ohta, A. (2000). Rethinking recasts: A learner-centered examination of corrective feedback in the Japanese classroom. In J.K. Hall & L. Verplaester (Eds.), *The construction of second and foreign language learning through classroom interaction* (pp. 47-71). Mahwah, NJ: Erlbaum.

Palmeri, T. (1997). Exemplar similarity and the development of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 324-354.

Panova, I., & Lyster, R. (2002). Patterns of corrective feedback and uptake in an adult ESL classroom. *TESOL Quarterly, 36,* 573-595.

Pawley, A. & Syder, F. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J.C. Richards & R. Schmidt (Eds.), *Language and communication.* (pp. 119-225), London: Longman.

Pica, T., Young, R. and Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly, 21*, 737-758.

Philp, J. (2003). Constraints on "noticing the gap": Non-native speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition, 25*, 99-126.

Pinker, S. (1999). *Words and Rules: The Ingredients of Language.* New York: Basic Books.

Pinker, S. & Prince, A. (1994). Regular and irregular morphology and the psychological status of rules of grammar. In S. Lima, R. Corrigan, & G. Iverson (Eds.), *The reality of linguistic rules* (pp. 353-388). Amsterdam: Benjamins.

Pinker, S.& Ullman, M. (2002). The past and future of the past tense. *Trends in Cognitive Science, 6*, 456-463.

Quirk, R. (1972). *A grammar of contemporary English.* London: Longman.

Rao, Z. (2002). Chinese students' perceptions of communicative and non-communicative activities in the EFL classroom. *System, 30*, 85-105.

Ran, Z. (2007). *Individual differences and cultural factors in English learning strategies*. Shanghai: Shanghai Foreign Language Education Press.

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

Robinson, P. (1995). Review article: Attention, memory and the "noticing" hypothesis. *Language Learning, 45*, 283-331.

Robinson, P. (2003). Attention and memory in SLA. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 631-678). Oxford: Blackwell.

Robinson, P. & Ha, M. (1993). Instance theory and second language rule learning under explicit conditions. *Studies in Second Language Acquisition, 13*, 413-438.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for second language acquisition: A meta-analysis of the research. In J. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp. 133-163). Amsterdam: John Benjamins.

Salaberry, M. (2000). The acquisitioin of English past tense in an instructional setting. *System, 28*, 135-152.

Segalowitz, N. (1997). Individual differences in second language acquisition. In A. de Groot &J. Kroll (Eds.), *Tutorials in bilingualism: psycholinguistic perspectives* (pp. 85-112). Mahwah, NJ: Lawrence Erlbaum.

Segalowitz, N. (2002). Automaticity and second languages. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 382-409). Oxford: Blackwell.

Seliger, H. (1983). The language learner as linguist: Of metaphors and realities. *Applied Linguistics, 4*, 179-191.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, *11*, 129-158.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition, 14*, 357-385.

Schmidt, R. (1995). *Attention and awareness in foreign language learning*. Honolulu, HI: University of Hawaii Press.

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

Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. In R. Day (Ed.), *Talking to learn: Conversation in second language acquisition* (pp. 237-326). Rowley, MA: Newbury House.

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

Seedhouse, P. (1997). The case of the missing "no": The relationship between pedagogy and interaction. *Language Learning, 47*, 547-583.

Seedhouse, P. (2004). *The interactional architecture of the language classroom: A conversation analysis perspective*. Malden, MA: Blackwell.

Sharwood Smith, M. (1986). The competence/control model, cross-linguistic influence and the creation of new grammars. In E. Kellerman & M. Sharwood Smith (Eds.), *Crosslinguistic influence in second language acquisition* (pp. 10-20). Oxford: Pergamon Press.

Sheen, Y. (2004). Corrective feedback and learner uptake in communicative classrooms across instructional settings. *Language Teaching Research, 8*, 263-300.

Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and learner uptake. *Language Teaching Research, 10*, 361-392.

Sheen, Y. (2007). The effects of corrective feedback, language aptitude and learner attitudes on the acquisition of English articles. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 301-322). Oxford: Oxford University Press.

Shirai, Y. (1991). Primacy of aspect in language acquisition: Simplified input and prototype. Unpublished Doctoral Dissertation, University of California, Los Angeles.

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Skehan, P. (1996). A framework for the implementation of task based instruction. *Applied Linguistics, 17*, 38-62.

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

Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching*, *29*, 1-15.

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

Spada, N. & Fröhlich, M. (1995). *COLT. Communicative Orientation of Language Teaching observation scheme: Coding conventions and applications*. Sydney, Australia: National Centre for English Language Teaching and Research.

Stern, H. (1990). Analysis and experience as variables in second language pedagogy. In B. Harley, P. Allen, J. Cummins, & M. Swain (Eds.), *The development of second language proficiency* (pp. 93-109). New York: Cambridge University Press.

Stern, H. (1992). *Issues and options in language teaching.* Oxford: Oxford University Press.

Suzuki, M. (2004, March). Corrective feedback and learner uptake in adult ESL classrooms. *Working Papers in TESOL and Applied Linguistics*. *4*, Retrieved April 13, 2004, from http://journals.tc-library.org/index.php/tesol/article/view/58

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Rowley, MA: Newbury House.

Swain, M. (1993). The output hypothesis: Just speaking and writing aren't enough. *The Canadian Modern Language Review, 50*, 158-164.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principles and practice in the study of language* (pp. 125-144). Oxford: Oxford University Press.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64-81). New York: Cambridge University Press.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Modern Language Journal, 82*, 320-37.

Swain, M. & Lapkin, S. (2001). Focus on form through collaborative dialogue: exploring task effects. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: second language learning, teaching, and testing* (pp. 99-118). New York: Longman.

Swain, M., & Lapkin, S. (2002). Talking it through: two French immersion learners' response to reformulation. *International Journal of Educational Research, 37*, 285-304.

Tomasello, M., & Herron, C. (1988). Down the garden path: Inducing and correcting overgeneralization errors in the foreign language classroom. *Applied Psycholinguistics, 9*, 237-246.

Tomasello, M., & Herron, C. (1989). Feedback for language transfer errors. *Studies in Second Language Acquisition, 11*, 385-395.

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

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

Truscott, J. (1999). What's wrong with oral grammar correction. *The Canadian Modern Language Review, 55*, 437-456.

Tsang, W. (2004). Feedback and uptake in teacher-student interaction: An analysis of 18 English lessons in Hong Kong secondary classrooms. *Regional Language Centre Journal, 35*, 187-209.

White, L. (1987). Against comprehensible input: The input hypothesis and the development of L2 competence. *Applied Linguistics, 8*, 95-110.

William, J. (2001). The effectiveness of spontaneous attention to form. *System, 29*, 325-340.

Wolfram, W. (1985). Variability in tense marking: a case for the obvious. *Language Learning, 35*, 229-253.

Yang, Y. (2006). Interactional feedback and uptake in Chinese EFL classrooms. Paper presented at the joint conference of AAAL/CAAL, Montreal, June, 2006.

Yu, L. (2001). Communicative language teaching in China: Progress and resistance. *TESOL Quarterly, 35*, 194-198.

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

McGill

Research Ethics Board Office
McGill University
845 Sherbrooke Street West
James Administration Bldg., rm 419
Montreal, QC H3A 2T5

Tel: (514) 398-6831
Fax: (514) 398-4644
Ethics website: www.mcgill.ca/researchoffice/compliance/human/

Research Ethics Board II
Certificate of Ethical Acceptability of Research Involving Humans

REB File #: 29-0607

Project Title: Recasts, prompts and Chinese EFL learners' acquisition of English past tense

Principal Investigator: Yingli Yang **Department:** DISE, Faculty of Education

Status: Ph.D.s student **Supervisor:** Prof. R. Lyster

Funding Agency and Title (if applicable): McGill SSHRC Internal

This project was reviewed on July 11, 2007 by

Expedited Review ✓
Full Review ____

Mark Baldwin, Ph.D.
Chair, REB II

Approval Period: July 16, 2007 **to** July 15, 2008

This project was reviewed and approved in accordance with the requirements of the McGill University Policy on the Ethical Conduct of Research Involving Human Subjects and with the Tri-Council Policy Statement: Ethical Conduct For Research Involving Humans.

---

* All research involving human subjects requires review on an annual basis. A Request for Renewal form should be submitted at least one month before the above expiry date.
* When a project has been completed or terminated a Final Report form must be submitted.
* Should any modification or other unanticipated development occur before the next required review, the REB must be informed and any modification can't be initiated until approval is received.

Appendix A

INFORMED CONSENT FORM FOR SCHOOL ADMINISTRATORS

Dear Sir/Madam:

I am writing to ask for your permission for my research in your school in Sep, 2007.

I am a Ph.D candidate in Faculty of Education, McGill University under the supervision of Dr. Roy Lyster. I am going to conduct a study on classroom interaction in an EFL (English as a Foreign Language) classroom in China. It is hoped that the study will contribute to the understanding of the role of error correction and teacher-student interaction in second language education in classroom settings.

If permitted, the researcher will pay three visits to a few classrooms and select 6 classes of the same proficiency level for the study. The study will take place in June, 2007. For research purposes, the students will be asked to participate in 4 communicative activities (each lasting about 30 minutes) that will be incorporated into their speaking practices. Three tests as well as a short questionnaire will be administered to trace students' second language development. The oral tests and treatment sessions will be audio-recorded with the researcher's presence. The entire study will take approximately one month. The researcher will meet with the teachers to discuss timelines that will be best commensurate with their schedules.

Participation in this research is totally voluntary. Students and teachers who choose not to participate in the study or to be taped will not be penalized in any respect. Furthermore, the students who do not wish to participate will be provided with activities and tests used in the study as exercise. The results of the study will be used for research purpose only, accessible only to the researcher, and will not count towards the students' grade or any other evaluation of either the students or the teachers. The audio-recordings will be disposed once the analysis is completed. Moreover, the results will be kept completely confidential; the names and any other personal information of the school, students, or teachers will not be used in any report of the study. At any time of the research, the students and the teachers can withdraw from the study without any negative consequences.

Should you agree with the research plan, please sign the consent form at the bottom of this letter. Should you have any questions or hesitations about the research, please contact me by e-mail or by phone. I greatly appreciate your consideration of my request and your support in the research.

Sincerely,
Yingli Yang

Ph.D Candidate
McGill University

E-mail: yingli.yang@mail.mcgill.ca
Tel: 0431-85674602

Consent Form

I have read the above and I understand the purpose and procedures of the study. I hereby agree with the researcher's plan of the study and will give her the permission to do research in our school. I also understand that it is totally voluntary for the students and teachers to participate in the study and they may choose not to participate in the study or withdraw from the study without any negative consequences.

Name (please print) _____

Signature _____ Date _____

I hereby agree with the researcher's plan of the study and will give her the permission to do audio-record the oral tests and treatment sessions of the selected classes in our school .I also understand that the results of the study are kept completely confidential and accessible to the researcher only.

Name (please print) _____

Signature _____ Date _____

Instructions for Teachers on Feedback Types

Recast Group:
If the student makes an error on the use of past tense, you should provide them with the correct form in a natural way so that the form becomes part of a correct utterance.
For example,
Student: I make a cake yesterday.
Teacher:
A: Oh, you made a cake. That's great. What else did you make?
or
B: You made a cake. Why did you make a cake?
C: Wonderful! What happened after you made the cake?

The key idea is that you correct the error, and at the same time, maintain the natural flow of the conversation.

Prompt Group:
If the student makes an error on the use of past tense, you should use the following techniques to elicit the correct forms from the students, allow them to say the correct forms themselves instead of giving them the correct forms.
For example,
Student: I make a cake yesterday.
Teacher
A: You make a cake yesterday?
Or
B: Do we say "make" for a past event?
C: Use past tense.
D: We don't say "make". What tense do we use for a past event?

The idea is that you withhold the correct form and use various cues to allow the students to correct their own errors.

Control Group

When the student makes an error in the communicative activity, you can temporarily ignore the errors and continue with the activities. At the end of the activity, you provide the students with the list of regular and irregular past tense forms and ask the students to reflect on them.

Appendix B

Example of Treatment Activities

1. Dictogloss Task

Divide the students into pairs. Read this passage twice at a normal speed and ask the students to take down notes, then ask the students to compare their text with their partner and write up a complete story. Ask students in each pair to retell part of the story.

It **cast** a gloom over the boat, there being no mustard. We **ate** our beef in silence. Existence **seemed** hollow and uninteresting. We **thought** of the happy days of childhood, and **sighed**. We **brightened** up a bit, however, when George **drew** out a tin of pine-apple, and **rolled** it into the middle of the boat, we **felt** that life **was** worth living after all.
We are very fond of pine-apples, all three of us. We **looked** at the picture on the tin; we **thought** of the juice. We **smiled** at one another, and Harris **got** a spoon ready.
Then we **looked** for the knife to open the tin with. We **turned** out everything in the hamper. We **turned** out the bags. We **pulled** up the boards at the bottom of the boat. We **took** everything out on to the bank and **shook** it. There **was** no tin-opener to be found.
Then Harris **tried** to open the tin with a pocket-knife, and **broke** the blade and **cut** himself badly; and George **tried** a pair of scissors, and the scissors **flew** up, and nearly **put** his eye out. And the tin **rolled** over, uninjured, and **broke** a tea cup.Then we all **got** mad. We **took** that tin out on the bank, and Harris **went** up into a field and **got** a big sharp stone. George **held** the sharp end of his stone against the top of it. I **went** back into the boat and **brought** out the mast, **gathered** up all my strength and **brought** it down.
It **was** George's straw hat that **saved** his life that day.... Harris **got** off with merely a flesh wound....
As for the tin...
There **was** one great dent across the top that **had** the appearance of a mocking grin, and it **drove** us crazy, so that Harris **caught** it up, and **threw** it far out into the middle of the river, and as it **sank**, we **got** into the boat and **rowed** away from the spot...

Adapted from

http://www.englishforjapanese.com/exercises/verb%20forms/15%20past%20tense%20review.html

2. Question-and-answer

Question-and-answer (Teacher's guide)

Please ask the students to answer aloud using "Yes.... " and incorporate the information in their card to complete the sentence. Conjugate the verb into the proper tense.

01. Did the concert begin on time?

01. at 6 pm.

02. Did Mr Jones lose his temper?

02. because of his son

03. Did Mrs Jones travel last year?

03. a lot with her husband

04. Did they arrive here early?

04. half an hour before dinner (begin)

05. Did they drink tea every day?

05. two cups of tea when she (live) in England

Question and answer-Student's guide

Instruction: Please answer aloud using "Yes.... " and incorporate the information in your card to complete the sentence.

01. at 6 pm.

02. because of his son

03. a lot with her husband

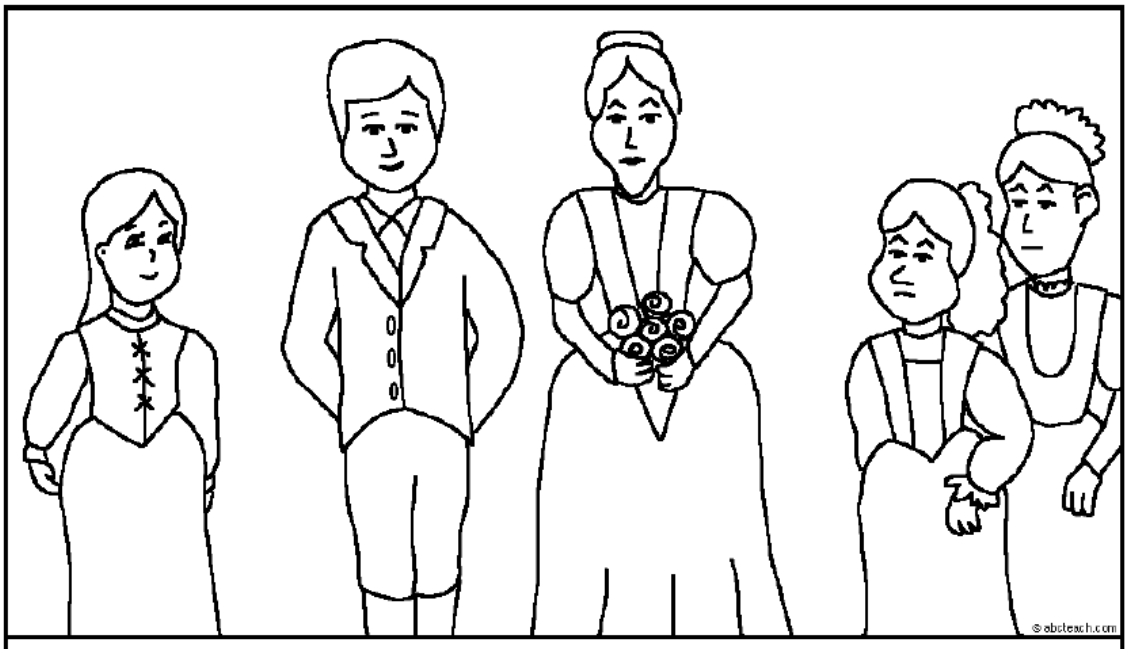
04. half an hour before dinner (begin)

05. two cups of tea when she (live) in England

3. Picture-cued narrative

Divide the students into groups of 3 or 4. Each group describes one picture and each student in the pair should say at least one sentence.

Examples of the pictures



Adapted from: http: www.abcteach.comfreeccinderellasequencecards.pdf.

Procedures of the Activities (Control group)

1. Treatment using dictogloss tasks:

a). The teacher organizes the students in pairs and inform them of the procedures of the dictogloss.

b). The teacher reads the story in normal speed and the students take down the text.

c). Students compare their texts and discuss with their partners to write up a complete story.

d). Each pair presents their part of the story to the entire class. Each member reads one or two sentences and passes on to his/her partner. The does not provide feedback.

e). The teacher hands out the original text for students to compare with their own text.

2. Treatment using "question and answer" tasks:

a). The teacher gives instruction of the task procedure and hands out students' cards to each students.

b). The teacher calls on students to answer the questions by card numbers. The students answer the teacher's questions. The teacher can also scramble the cards and reassign them to the students after the first round and practice again. During this process, the teacher does not provide feedback to students' errors.

c). The teacher hands out a list of regular and irregular verbs and past tense forms to the students and ask them to reflect on their mistakes themselves.

3. Treatment using “Cinderella” Sequence Card

a) The teacher hands out scrambled pictures to the student in groups and gives instructions to the students.

b) Students work in groups and think about sentences that describe the content of the picture.

c) Students in each group come together to discuss the general sequence of the story.

d) The teacher calls on students to describe the pictures in sequence to finish the complete story. Each group describes one picture and each student in the pair should say at least one sentence. The teacher provides assistance only on difficult vocabulary, but not on grammar.

e) At the end of the activity, the teacher describes the picture in correct sequence him/herself in simple past tense.

Questionnaires

1. Background Information Questionnaire

This purpose of this questionnaire is to gather information about your background in learning English as well as bio-data. Please answer as completely as you can.

Background Information

1. Name: _____
2. Age: ____
3. Program/ Year of study: _____

4. Years of studying English:

5. Age of studying English in classroom settings:
   A. Before Primary school
   B. Since Primary school
   C. Since Junior middle school
   D. Since High school
   E. Since University

6. In which way do you practice oral English:
   A. With foreigners:
   B. With my foreign English teacher only
   C. With my Chinese English teacher
   D. With my classmates
   E. With strangers in the English corner
   F. Others (Please specify)

7. Which aspect do you think is the most important in oral English?
   A. Pronunciation
   B. Vocabulary
   C. Grammar
   D. Use of idiomatic expressions
   E. Clarity of meaning
   F. Fluency

2. Exit Questionnaire

1. Now that you have completed the tasks and the tests, what do you think they were all about?

A. They were practicing and testing writing.

B. They were practicing and testing my grammar.

C. They were practicing and testing listening and speaking.

D. They were practicing and testing my vocabulary.

2. Please write a short paragraph saying what you think you learned from these communicative tasks and tests.

3. Do you like the tasks you performed in class? Why do you like or dislike them?

4. Do you like to be corrected while you are speaking in class? Why or why not?

Appendix C

Testing Material

1. Oral Test (Version A)

Please read this story silently for three minutes. Try to remember the content of the story. After 3 minutes, this text will be removed and you are required to retell the story with the help of some word prompts.

A Crazy Beach Party

I flew to Korea for the first time in 2000. One of the first places I visited was a beach called Sunset Beach. I met a lot of foreigners there, as well as a Korean man who insisted in telling people that his name was Ricky. We sat on the beach, drank beer, and laughed a lot.

Since it was winter, we began to get cold outside. We walked to a nearby store and bought some beer. We stood near a large fire and started to sing as we drank beer. We all felt very happy. Suddenly, Ricky's hair caught on fire. Luckily, someone put it out.

Some people jumped and ran to keep warm, other smoked cigarette. A drunk girl fell to the ground and cut her knee. She bled really badly, so Ricky and I sent her to a nearby hospital. The doctor said the wound was not serious and warned us not to drink too much.

The next morning, at 6 am., the sun rose. It was another chilly winter morning. We drove home after breakfast and it only took about 30 minutes to get back to Seoul. I never expected such a crazy beach party.

Adapted from ESL Quiz center, http://www.eslgo.com/quizzes/irregpast2.html

Word cues for the oral test

Fly to Korea, in 2000

Visit a beach, Sunset beach

Meet foreigners, a Korean man, insist in telling people

Sit on the beach, drink beer, laugh

Begin to get cold

Walk to a store, buy beer

Stand near a large fire, start to sing

Feel happy

Ricky’s hair, catch on fire

Put it out

Jump and run, smoke

A drunk girl, fall to the ground, cut her knee

Bleed badly, send her to hospital

The doctor, say the wound not serious, warn

The sun rise, at 6am.

Drive home

Take 30 minutes

Never expect such a crazy party

2. Written Test (Delayed post-test version)

Please rewrite the story “Cinderella” using the following verbs and phrases. You should use all of the verbs and phrases listed below and narrate the events in sequence. Use **SIMPLE PAST TENSE of these verbs ONLY** throughout the story (**Please do not use infinitive or participles or negation with these verbs**). However, you can add new verbs and use other tenses appropriately with the new verbs if necessary. Also try to connect sentences by using adverbs “first, then, consequently” and so on to write a coherent story.

Verbs to be used:

live, seem, sigh, share, murmur, mean,
throw, ink, fear, appear, turn,
gather, draw, smile, flee, find

The Scoring Procedure

Accuracy in the present study is the operationalized as "the correct use of past tense forms in appropriate past tense context". The general criteria for coding and scoring the oral data are a) the suppliance of the past tense forms in obligatory context a) the accuracy of the past tense forms of these English verbs.

Specifically, the combination of a) and b) (i.e. the use of the correct form of the simple past tense of this particular verb in appropriate context) would grant the student "1" on the scoring chart next to the verb. For example, if the student says "I flew to Korea for the first time in the year 2000", then, next to the verb "fly" on the scoring sheet, you can put "1" because he/she uses the correct past tense form of the verb "fly" (i.e. flew) in the appropriate past tense context. Similarly, as the student continues to narrate the story, the tense is already set in the past (in the year 2000). Therefore, the student is supposed to use past tense consistently throughout the text.

However, if the student fails to supply the correct form of the past tense or use other tenses (including other past tense or present tense), the scoring procedure is as follows.

You should score "o" next to the verb, if the student:

a) use other past tenses (e.g. past continuous tense, or past perfect tense, which then will not allow you to know whether he knows how to use simple past tense of this verb properly.)
b) completely miss the verb (according to the context, the student uses other verbs to replace this particular verb or simply ignore this part of the story and goes on with narrating the next sentence). However, sometimes, the students would go back and narrate this sentence again later. In this case, you will still score the verb according to these criteria.
c) Use of infinitive form of the verb (e.g. "We began **to jump** and run" instead of "We jumped and ran to keep warm") or past participle (e.g. Ricky's hair was caught on fire" instead of "Ricky's hair caught on fire").
d) use bare form of the verb in past tense context (e.g. "I fly to Korea in 2000").
e) use wrong/hybrid past tense form of this verb (e.g. I flied to Korea in 2000 or He was

bleed badly).

f) Use of present tense in obligatory past tense context (e.g. It takes me 30 minutes to go back to Seoul the next morning.)

Note: In certain circumstances, the students would say a wrong form first and then self-correct. This is still considered correct if the second form used is correct. So a mark of "1" will be given to that verb.

Coding Sheet

Class/name																							
fly																							
visit																							
meet																							
insist																							
was																							
walk																							
buy																							
sit																							
drink																							
laugh																							
stand																							
start																							
feel																							
fall																							
bleed																							
send																							
say																							
warn																							
was																							
begin																							
jump																							
smoke																							
catch																							
rise																							
drive																							
take																							
expect																							

Appendix D

Tables of Multiple Comparisons of Group Means

Table D1

Comparison of means of each group on the use of overall past tense over time in the oral tests

Unconstrained (Free Combinations) Step-Down Tests

Contrast	Estimate	Standard Error	Pr > |t| Raw	Bon	Adj	SE(AdjP)
recast, delay-post	-5.3924	3.6101	0.1372	0.4116	0.3572	0.000813
recast, delay-pre	7.5357	3.8214	0.0503	0.2013	0.1811	0.000546
recast, post-pre	12.9281	4.0714	0.0018	0.0108	0.0101	0.000097
prompt, delay-post	0.1612	3.7682	0.9659	0.9659	0.9659	0
prompt, delay-pre	14.4612	4.1012	0.0005	0.0044	0.0042	0.000055
prompt, post-pre	14.3000	4.1230	0.0007	0.0047	0.0045	0.000058
control, delay-post	-4.6426	3.3164	0.1635	0.4116	0.3572	0.000813
control, delay-pre	9.9844	3.8106	0.0096	0.0481	0.0456	0.000200
control, post-pre	14.6270	3.7717	0.0002	0.0014	0.0013	0.000019

Table D2

Comparison of means of different groups at each time point on the overall use of past tense in the oral tests

Unconstrained (Free Combinations) Step-Down Tests

Contrast	Estimate	Standard Error	Pr > |t| Raw	Bon	Adj	SE(AdjP)
delay, recast-prompt	-9.9289	5.3369	0.0646	0.5168	0.3243	0.00117
delay, recast-control	2.5883	5.1353	0.6149	1.0000	0.8588	0.00124
delay, prompt-control	12.5172	5.3292	0.0200	0.1801	0.1305	0.000713
post, recast-prompt	-4.3753	4.9907	0.3819	1.0000	0.8588	0.00124
post, recast-control	3.3381	4.7631	0.4844	1.0000	0.8588	0.00124
post, prompt-control	7.7134	4.7494	0.1063	0.7438	0.4364	0.00129
pre, recast-prompt	-3.0034	6.2600	0.6320	1.0000	0.8588	0.00124
pre, recast-control	5.0370	6.0569	0.4068	1.0000	0.8588	0.00124
pre, prompt-control	8.0404	6.2600	0.2008	1.0000	0.6300	0.00141

Table D3

Comparison of means of each group on the use of overall past tense over time in the written tests

Unconstrained (Free Combinations) Step-Down Tests

Contrast	Estimate	Standard Error	Pr > |t| Raw	Bon	Adj	SE(AdjP)
recast, delay-post	-4.7654	3.3182	0.1527	0.2742	0.2542	0.000540
recast, delay-pre	15.8868	4.5971	0.0007	0.0041	0.0039	0.000052
recast, post-pre	20.6522	4.0740	<.0001	<.0001	<.0001	0
prompt, delay-post	-6.4394	3.3302	0.0547	0.1642	0.1545	0.000391
prompt, delay-pre	26.1364	4.6554	<.0001	<.0001	<.0001	0
prompt, post-pre	32.5758	4.1656	<.0001	<.0001	<.0001	0
control, delay-post	-6.8363	3.1560	0.0316	0.1265	0.1161	0.000397
control, delay-pre	6.4494	4.3188	0.1371	0.2742	0.2542	0.000540
control, post-pre	13.2857	3.8646	0.0007	0.0041	0.0039	0.000052

Table D4

Comparison of means of different groups at each time point on the overall use of past tense in the written tests

Unconstrained (Free Combinations) Step-Down Tests

Contrast	Estimate	Standard Error	Pr > |t| Raw	Bon	Adj	SE(AdjP)
delay, recast-prompt	-3.9254	4.7693	0.4116	0.9005	0.6500	0.00108
delay, recast-control	8.4480	4.6155	0.0689	0.4820	0.3340	0.00108
delay, prompt-control	12.3735	4.6256	0.0082	0.0653	0.0566	0.000347
post, recast-prompt	-5.5995	4.4065	0.2055	0.9005	0.5568	0.00128
post, recast-control	6.3771	4.2594	0.1361	0.8166	0.5081	0.00130
post, prompt-control	11.9765	4.3098	0.0060	0.0543	0.0463	0.000332
pre, recast-prompt	6.3241	5.5950	0.2598	0.9005	0.5637	0.00119
pre, recast-control	-0.9894	5.3705	0.8540	0.9005	0.8540	0
pre, prompt-control	-7.3135	5.4349	0.1801	0.9005	0.5534	0.00129

Table D5

Comparison of means of each group on the use of irregular past tense over time in the oral tests

Unconstrained (Free Combinations) Step-Down Tests

Contrast	Estimate	Standard Error	Pr > |t| Raw	Bon	Adj	SE(AdjP)
recast, delay-post	-8.9514	3.7048	0.0168	0.0672	0.0648	0.000201
recast, delay-pre	6.3640	3.6566	0.0837	0.1674	0.1601	0.000348
recast, post-pre	15.3153	4.4353	0.0007	0.0051	0.0049	0.000063
prompt, delay-post	-1.1860	3.8441	0.7581	0.7581	0.7581	0
prompt, delay-pre	13.6393	3.9170	0.0006	0.0051	0.0049	0.000063
prompt, post-pre	14.8253	4.4870	0.0012	0.0070	0.0069	0.000043
control, delay-post	-7.1092	3.3658	0.0362	0.1086	0.1042	0.000273
control, delay-pre	9.8529	3.6408	0.0075	0.0377	0.0368	0.000127
control, post-pre	16.9621	4.1083	<.0001	0.0005	0.0005	0.000025

Table D6

Comparison of means of different groups at each time point on the use of irregular past tense in the oral tests

Unconstrained (Free Combinations) Step-Down Tests

Contrast	Estimate	Standard Error	Pr > |t| Raw	Bon	Adj	SE(AdjP)
delay, recast-prompt	-11.1339	5.5249	0.0455	0.3639	0.2438	0.00100
delay, recast-control	0.7333	5.3219	0.8906	1.0000	0.9410	0.000923
delay, prompt-control	11.8673	5.5145	0.0328	0.2955	0.1988	0.000924
post, recast-prompt	-3.3685	5.3457	0.5295	1.0000	0.9410	0.000923
post, recast-control	2.5755	5.1028	0.6144	1.0000	0.9410	0.000923
post, prompt-control	5.9440	5.0777	0.2434	1.0000	0.7095	0.00141
pre, recast-prompt	-3.8586	6.3774	0.5460	1.0000	0.9410	0.000923
pre, recast-control	4.2222	6.1705	0.4948	1.0000	0.9410	0.000923
pre, prompt-control	8.0808	6.3774	0.2069	1.0000	0.6763	0.00143

Table D7

Comparison of means of each group on the use of regular past tense over time in the oral tests

Unconstrained (Free Combinations) Step-Down Tests

Contrast	Estimate	Standard Error	Pr > |t| Raw	Bon	Adj	SE(AdjP)
recast, delay-post	0.0555	5.4705	0.9919	1.0000	0.9994	0.000103
recast, delay-pre	9.5072	5.6762	0.0959	0.4546	0.3371	0.000988
recast, post-pre	9.4517	5.3946	0.0817	0.4546	0.3371	0.000988
prompt, delay-post	2.8723	5.6926	0.6146	1.0000	0.9423	0.000878
prompt, delay-pre	16.6244	6.1186	0.0073	0.0658	0.0585	0.000305
prompt, post-pre	13.7522	5.4786	0.0131	0.1045	0.0912	0.000406
control, delay-post	0.1243	5.1109	0.9806	1.0000	0.9994	0.000103
control, delay-pre	10.1526	5.6800	0.0758	0.4546	0.3371	0.000988
control, post-pre	10.0283	5.0025	0.0467	0.3268	0.2505	0.000837

Table D8

Comparison of means of different groups at each time point on the use of regular past tense in the oral tests

Unconstrained (Free Combinations) Step-Down Tests

Contrast	Estimate	Standard Error	Pr > |t| Raw	Bon	Adj	SE(AdjP)
delay, recast-prompt	-8.4102	6.5553	0.2013	1.0000	0.7154	0.00142
delay, recast-control	6.0212	6.2849	0.3395	1.0000	0.7904	0.00134
delay, prompt-control	14.4314	6.5585	0.0292	0.2627	0.1969	0.000800
post, recast-prompt	-5.5934	5.9016	0.3447	1.0000	0.7904	0.00134
post, recast-control	6.0900	5.6026	0.2787	1.0000	0.7904	0.00134
post, prompt-control	11.6834	5.5454	0.0367	0.2934	0.2193	0.000837
pre, recast-prompt	-1.2929	7.4517	0.8625	1.0000	0.8625	0
pre, recast-control	6.6667	7.2100	0.3565	1.0000	0.7904	0.00134
pre, prompt-control	7.9596	7.4517	0.2870	1.0000	0.7904	0.00134

Table D9

Comparison of means of each group on the use of irregular past tense over time in the written tests

Unconstrained (Free Combinations) Step-Down Tests

Contrast	Estimate	Standard Error	Pr > |t| Raw	Bon	Adj	SE(AdjP)
recast, delay-post	-1.9909	5.6553	0.7252	1.0000	0.9633	0.000737
recast, delay-pre	23.2265	5.4641	<.0001	0.0002	0.0002	0.000017
recast, post-pre	25.2174	6.1955	<.0001	0.0004	0.0004	0.000018
prompt, delay-post	-1.8182	5.7074	0.7504	1.0000	0.9633	0.000737
prompt, delay-pre	36.3636	5.5094	<.0001	<.0001	<.0001	0
prompt, post-pre	38.1818	6.3347	<.0001	<.0001	<.0001	0
control, delay-post	-2.2957	5.3971	0.6711	1.0000	0.9633	0.000737
control, delay-pre	5.1812	5.1307	0.3139	1.0000	0.7693	0.00130
control, post-pre	7.4769	5.9041	0.2070	1.0000	0.6303	0.00131

Table D10

Comparison of means of different groups at each time point on the use of irregular past tense in the written tests

Unconstrained (Free Combinations) Step-Down Tests

Contrast	Estimate	Standard Error	Pr > |t| Raw	Bon	Adj	SE(AdjP)
delay, recast-prompt	-9.6194	6.8981	0.1649	0.8245	0.5354	0.00133
delay, recast-control	6.8078	6.6744	0.3091	0.9273	0.6510	0.00124
delay, prompt-control	16.4272	6.6906	0.0150	0.1353	0.1071	0.000573
post, recast-prompt	-9.4466	7.8187	0.2286	0.9142	0.6061	0.00131
post, recast-control	6.5031	7.5650	0.3911	0.9273	0.6510	0.00124
post, prompt-control	15.9497	7.6542	0.0386	0.2902	0.2134	0.000850
pre, recast-prompt	3.5178	7.2002	0.6257	0.9273	0.6510	0.00124
pre, recast-control	-11.2375	6.9114	0.1057	0.6343	0.4277	0.00124
pre, prompt-control	-14.7552	6.9942	0.0363	0.2902	0.2134	0.000850

Table D11

Comparison of means of each group on the use of regular past tense over time in the written tests

Unconstrained (Free Combinations) Step-Down Tests

Contrast	Estimate	Standard Error	Pr > |t| Raw	Bon	Adj	SE(AdjP)
recast, delay-post	-6.7501	3.5730	0.0605	0.1701	0.1535	0.000490
recast, delay-pre	10.6412	5.5485	0.0567	0.1701	0.1535	0.000490
recast, post-pre	17.3913	4.9652	0.0006	0.0041	0.0038	0.000054
prompt, delay-post	-9.7403	3.5845	0.0072	0.0289	0.0280	0.000133
prompt, delay-pre	18.8312	5.6292	0.0010	0.0060	0.0059	0.000043
prompt, post-pre	28.5714	5.0768	<.0001	<.0001	<.0001	0
control, delay-post	-10.2328	3.3842	0.0029	0.0143	0.0139	0.000090
control, delay-pre	7.2686	5.2138	0.1650	0.1701	0.1650	0
control, post-pre	17.5014	4.6918	0.0003	0.0020	0.0019	0.000050

Table D12

Comparison of means of different groups at each time point on the use of regular past tense in the written tests

Unconstrained (Free Combinations) Step-Down Tests

Contrast	Estimate	Standard Error	Pr > |t| Raw	Bon	Adj	SE(AdjP)
delay, recast-prompt	0.1387	4.7417	0.9767	1.0000	0.9767	0
delay, recast-control	9.7032	4.5961	0.0361	0.2891	0.2233	0.000806
delay, prompt-control	9.5645	4.5975	0.0389	0.2891	0.2233	0.000806
post, recast-prompt	-2.8515	3.5181	0.4187	1.0000	0.7949	0.00123
post, recast-control	6.2205	3.4070	0.0695	0.4173	0.3140	0.000956
post, prompt-control	9.0720	3.4471	0.0092	0.0831	0.0717	0.000379
pre, recast-prompt	8.3286	6.8096	0.2229	1.0000	0.6546	0.00132
pre, recast-control	6.3306	6.5364	0.3341	1.0000	0.7826	0.00129
pre, prompt-control	-1.9980	6.6147	0.7630	1.0000	0.9424	0.000817