The effectiveness of enhanced and unenhanced recasts on secondary school students' past tense usage in Hong Kong

Volume I—Thesis

Ching Ching Lai (Florence)

Institute of Education, University of London

Submitted for the degree of Doctor of Philosophy

December 2011
Abstract

Hong Kong secondary school students often lack the procedural knowledge of using English grammar in speaking, though they possess the relevant declarative knowledge from their education. Students’ tenseless L1 and lack of L2 exposure could have been the reasons. The present study therefore investigated if individualized feedback on students’ use of past tense in their spoken narratives would facilitate their development. Recasts were chosen as the target feedback, because they intertwine with the ongoing speaking flow and provide model reformulation to ease learners’ online cognitive load. However, controversies over the effectiveness of recasts lie on their variable explicitness and implicitness. Moreover, research to date has not sufficiently examined both the explicit and implicit implementation of recasts in the context of Hong Kong.

As a result, the present study explored the immediate and overtime effectiveness of implicit and explicit recasts over no feedback on Hong Kong learners’ use of past tense in their spoken narratives elicited from cartoon strips. The comparative effectiveness of the two recasts was also probed. Investigation into some students’ perception towards the feedbacks was conducted through stimulated recalls, to suggest causes behind the different effectiveness of recasts.

The study found that both recasts were more effective than no feedback on past tense in the short term. Opposite to what was hypothesized, only the effectiveness of implicit recasts sustained overtime; and both recasts were not significantly different from each other at any time. From students’ recalls of their thoughts, task demands, the online speaking mode, students’ online cognitive constraints, students’ deficient L2 mastery, and the different effects of recasts and
students' uptakes may have been the underlying reasons prompting the variable effectiveness of recasts.

Overall, the present study suggested that recasts in any degrees of explicitness can be effective or not, depending on how controlled they are in their implementation. Explicit recasts may not be sustainable, possibly due to their negative effect to learners. The suitability of recasts for students' online cognitive capacity and style may affect the explicitness and effectiveness of recasts. The online demand of a speaking task, students' L2 readiness for a specific task, and learners' different perception towards the function and their uptake of recasts may also affect the effectiveness of recasts.
Acknowledgements

I would like to express my heartfelt gratitude to people who have supported the accomplishment of this thesis in various ways.

My supervisor, Professor Richard Andrews, has been motivating me to move forward with his continuous encouragement and advice. He has given me freedom in expressing my thoughts and developing arguments, together with his helpful guidance. His positive and constructive comments provided the balance between upholding my own ideas and integrating his suggestions. His prompt responses to all my questions and requests were especially considerate.

My upgrade committee members, Dr. Amos Paran and Dr. Sian Preece, have given me green light to continue my PhD thesis by acknowledging my ability to proceed from an MPhil student status to a doctorate one. Their approvals have also raised my confidence in making the thesis a mission possible. Most importantly, their, as well as my internal reader Dr. John O’Regan’s, advice on the different aspects of my thesis has introduced me to the different angles of improvement. I am also thankful for people (e.g. Dr. Catherine Walter, Professor Norbert Pachler, Professor Gunther Kress, and Dr. Richard Freeman) who have contributed their advice at the different stages of my PhD studies. Last but not least, my examiners Professor David Block and Dr. Andrea Revesz have enlightened me with their critical perspectives.

The study would not have been possible without the help of the principal of the target school. She not only gave me the access to her school, but also the full support and cooperation of her teachers and students in making the entire data collection process smooth and successful. The teachers and students were kind enough to sacrifice their lesson time and holidays respectively to participate in the study.
I am greatly indebted to my parents and husband for unconditionally supporting me emotionally and financially throughout this hard but rewarding journey. Their love has offered me a total peace of mind and concentration on my PhD studies. My husband and his friend, Joffee Lam, have also assisted me in dealing with the technical and statistical aspects of my data analyses. In daily living, my mother-in-law has played an unforgettable role by feeding me with nutritious food during my thesis-writing stage and pregnancy. Speaking of which, I must also give credit to the little precious boy inside me, who has treated me well and given me strength to write my thesis.
# TABLE OF CONTENTS

**Volume I: Thesis**

Chapter 1 Introduction ................................................................................................... 1  
1.1 Thesis Topic and Development ............................................................................ 1 
1.2 Thesis Organization .............................................................................................. 6 

Chapter 2 Literature Review I - Theoretical Variables of Recasts ................................ 9  
2.1 Introduction .......................................................................................................... 9 
2.2 Input, Output and Interaction Hypotheses .......................................................... 14 
2.3 Corrective Feedback Classroom Studies ............................................................ 20 
2.4 Noticing .............................................................................................................. 30 
2.5 Response to Recasts ........................................................................................... 36 
2.6 Prompts and Recasts .......................................................................................... 40 
2.7 Frequency and Saliency ..................................................................................... 54 
2.8 Target Form Influence ....................................................................................... 64 
2.8.1 Interpretation of Form in Recasts ................................................................ 64 
2.8.2 Form Complexity / Difficulty ...................................................................... 68 
2.9 Conclusion .......................................................................................................... 76 

Chapter 3 Literature Review II - Methodological and EFL Variables of Recasts ...... 78 
3.1 Introduction ........................................................................................................ 78 
3.2 Test Influence ..................................................................................................... 78 
3.3 Role and Optimal Use of Feedback .................................................................... 83 
3.4 Learners’ Perception .......................................................................................... 91 
3.5 Learners’ Level/Developmental Readiness ...................................................... 97 
3.6 Task .................................................................................................................. 102 
3.6.1 Differences between the Two Attention and Task Models ......................... 103 
3.6.2 Measure of Speech and its L2 and Task Correlates ................................... 108 
3.6.3 Similarity between the Two Attention and Task Models ........................... 113 
3.6.4 Task and Recasts ........................................................................................ 115 
3.7 Modality ........................................................................................................... 120 
3.7.1 Effect of Modality on Learning of Form ................................................... 121 
3.7.2 Effect of Modality on Task Planning ......................................................... 125 
3.8 Context / Setting ............................................................................................... 128 
3.9 Conclusion ........................................................................................................ 140 
3.10 Research Gaps and Questions ........................................................................ 141 
3.11 Conclusion of Chapters 2 & 3 ....................................................................... 145 

Chapter 4 Methodology ............................................................................................. 147 
4.1 Introduction ...................................................................................................... 147 
4.2 Theoretical and Empirical Justification of the Research Design ....................... 148 
4.2.1 Quantitative Research Design .................................................................... 148 
4.2.2 Qualitative Research Design ...................................................................... 159 
4.3 The Pilot Study .................................................................................................. 172 
4.4 The Study Context in Hong Kong .................................................................... 174 
4.4.1 English Proficiency Level .......................................................................... 174 
4.4.2 Medium of Instruction .............................................................................. 176 
4.4.3 English Language Curriculum ................................................................... 177 
4.4.4 School Teachers ....................................................................................... 179
LIST OF FIGURES

Figure 1 — Ratios of the Past Tense Used in the Three Tests by Groups ............... 244
Figure 2 — Graphical Results of Two-way Mixed ANOVA on the Outcome Measure
   ‘Developing Past Form’ ........................................................................................................ 270
Figure 3 — Graphical Results of Two-way Mixed ANOVA on the Outcome Measure
   ‘Irregular Past Form’ ........................................................................................................... 272
Figure 4 — Graphical Results of Two-way Mixed ANOVA on the Outcome Measure
   ‘Non-Past Form’ ................................................................................................................. 274
Figure 5 — Graphical Results of Two-way Mixed ANOVA on the Outcome Measure
   ‘Modal Verbs’ .................................................................................................................... 276
Figure 6 — Graphical Results of Two-way Mixed ANOVA on the Outcome Measure
   ‘Same Past Form’ .............................................................................................................. 278
Figure 7 — Graphical Results of Two-way Mixed ANOVA on the Outcome Measure
   ‘Verbs in Monologue or Dialogue’ .................................................................................... 280
Figure 8 — Graphical Results of Two-way Mixed ANOVA on the Outcome Measure
   ‘Verbless L1 Transfer’ ......................................................................................................... 282
Figure 9 — Graphical Results of Two-way Mixed ANOVA on the Outcome Measure
   ‘No Vowel Change Past Form’ ........................................................................................ 284
Figure 10 — Graphical Results of Two-way Mixed ANOVA on the Outcome Measure
   ‘Regular Past Form’ .......................................................................................................... 286
Figure 11 — Distribution of ‘Past Tense Used Ratio’ of the Three Treatment Groups 290
Figure 12 — Distribution of ‘Errors’, ‘Recast Attempts’, ‘Recasts’, ‘Success Uptake
   Reformulation’ of NR and CR Groups .................................................................................. 293
Figure 13 — Distribution of ‘Unrecast Errors’, ‘Uptake No Reformulation’, ‘No
   Uptake’ and ‘Incorrect Uptake’ of NR and CR Groups ......................................................... 294
Figure 14 — Distribution of ‘Success Uptake Halfway’ of CR Group ......................... 295
LIST OF TABLES

Table 1—3 x 2 Factorial Research Design .................................................. 153
Table 2—Counterbalancing the Different Story Narrative Parts .................. 227
Table 3—Results of Two-way Mixed ANOVA on the Outcome Measure ‘Past Tense Used Ratio’ ................................................................. 245
Table 4—Results of One-way ANOVA and Post Hoc Multiple Comparisons among the Three Treatment Groups on ‘Past Tense Used Ratio’ at Each Time Point ............... 249
Table 5a—Results of Two-way Independent ANOVA of ‘Past Tense Used Ratio’ in Immediate Posttest ........................................................................... 253
Table 5b—Results of Two-way Independent ANOVA of ‘Past Tense Used Ratio’ in Delayed Posttest ............................................................... 253
Table 6—Categories of Verb Patterns ............................................................. 255
Table 7—Grouping 1: List of Verb Tokens with No Auxiliary ......................... 256
Table 8—Grouping 2: List of Verb Tokens Starting with ‘are’, ‘aren’t’, or ‘are not’ 258
Table 9—Grouping 3: List of Verb Tokens Starting with ‘can’, ‘can’t’, or ‘cannot’ 258
Table 10—Grouping 4: List of Verb Tokens Starting with ‘could’, ‘couldn’t’, or ‘could not’ ......................................................................................... 259
Table 11—Grouping 5: List of Verb Tokens Starting with ‘did’, ‘didn’t’, or ‘did not’ ......................................................................................... 259
Table 12—Grouping 6: List of Verb Tokens Starting with ‘do’, ‘don’t’, or ‘do not’ 260
Table 13—Grouping 7: List of Verb Tokens Starting with ‘had’, ‘hadn’t’, or ‘had not’ ......................................................................................... 260
Table 14—Grouping 8: List of Verb Tokens Starting with ‘has’, ‘hasn’t’, or ‘has not’ ......................................................................................... 260
Table 15—Grouping 9: List of Verb Tokens Starting with ‘is’, ‘isn’t’, or ‘is not’ .... 261
Table 16—Grouping 10: List of Verb Tokens Starting with ‘make’ ..................... 261
Table 17—Grouping 11: List of Verb Tokens Starting with ‘may’ or ‘may not’ ..... 261
Table 18—Grouping 12: List of Verb Tokens Starting with ‘must’, ‘mustn’t’, or ‘must not’ ......................................................................................... 262
Table 19—Grouping 13: List of Verb Tokens Starting with ‘see’ ....................... 262
Table 20—Grouping 14: List of Verb Tokens Starting with ‘should’, ‘shouldn’t’, or ‘should not’ ......................................................................................... 262
Table 21—Grouping 15: List of Verb Tokens Starting with ‘will’, ‘won’t’, or ‘will not’ ......................................................................................... 262
Table 22—Grouping 16: List of Verb Tokens Starting with ‘would’, ‘wouldn’t’, or ‘would not’ ......................................................................................... 263
Table 23—Grouping 17: List of Verb Tokens Starting with ‘were’, ‘weren’t’, or ‘were not’ ......................................................................................... 263
Table 24—Grouping 18: List of Verb Tokens Starting with ‘does’, ‘doesn’t’, or ‘does not’ ......................................................................................... 263
Table 25—Grouping 19: List of Verb Tokens Starting with ‘have’, ‘haven’t’, or ‘have not’ ......................................................................................... 264
Table 26—Grouping 20: List of Verb Tokens Starting with ‘was’, ‘wasn’t’, or ‘was not’ ......................................................................................... 264
Table 27—Grouping 21: List of Verb Tokens Starting with ‘saw’ ....................... 265
Table 28—Results of Two-way Mixed ANOVA on the Outcome Measure ‘Developing Past Form’ ........................................................................ 269
Table 29—Results of Two-way Mixed ANOVA on the Outcome Measure ‘Irregular Past Form’ ........................................................................ 271
Table 30 – Results of Two-way Mixed ANOVA on the Outcome Measure ‘Non-Past Form’ ................................................................. 273
Table 31 – Results of Two-way Mixed ANOVA on the Outcome Measure ‘Modal Verbs’ ................................................................. 275
Table 32 – Results of Two-way Mixed ANOVA on the Outcome Measure ‘Same Past Form’ ......................................................... 277
Table 33 – Results of Two-way Mixed ANOVA on the Outcome Measure ‘Verbs in Monologue or Dialogue’ .................................. 279
Table 34 – Results of Two-way Mixed ANOVA on the Outcome Measure ‘Verbless L1 Transfer’ ................................................. 281
Table 35 – Results of Two-way Mixed ANOVA on the Outcome Measure ‘No Vowel Change Past Form’ ........................................ 283
Table 36 – Results of Two-way Mixed ANOVA on the Outcome Measure ‘Regular Past Form’ .......................................................... 285
Table 37 – Results of One-way ANOVA and Post Hoc Multiple Comparisons among Three Feedback Treatments on Students’ Past Tense Used Ratio ......................................................... 290
Table 38—Mann-Whitney U Test on Variables across the Two Feedback Groups ................................................................. 291
Table 39—Identity Code for Learners in the Stimulated Recall Subset ....................................................................................... 298
Table 40—Transcription Code in Stimulated Recall Excerpts .................................................................................... 299
Chapter 1 Introduction

1.1 Thesis Topic and Development

The topic of this thesis is recasts. My first experience with the term *recasts* was in a Master’s introductory course on second language acquisition in New York. Recasts were briefly presented as one of the corrective feedback types used by ESL teachers. My hands-on experience with recasts was when I actually used recasts to give feedback to an ESL student in a teaching practicum. The student was assigned to an upper-intermediate ESL class, and was taught grammar through communicative activities. I recast his/her grammar mistake during the interview about his/her group discussion outcome. I chose to use recasts because they were convenient to use without interrupting the student’s speech too much. Regardless of the previous recast, the student kept making the same mistake. Moreover, s/he replied with a “huh?” after each following repeated recast. From the unsuccessful uptakes of the student, I began to have interest in what makes recasts effective.

Because of their intertwining with the ongoing interaction, recasts were introduced as an appropriate feedback type to be given during communication in the literature as well as in the course. However, my practicum experience indicated that recasts may not be effective as expected. Although recasts may not be effective in reality, they are often the first few feedback types that teachers would think of when correcting students’ errors during their interaction. Recasts seem theoretically sound in an interactive context, which leads to the question of whether recasts are also effective in a non-interactive context. Hong Kong, my hometown, is a typical EFL context which uses teacher-fronted classroom instruction to mainly deliver target language grammar. With no support of an interactive context plus no convention of
giving one-on-one corrective feedback in class, recasts may quickly be given a verdict of ineffectiveness and inappropriateness in the context of Hong Kong.

Recasts are still worth studying in the context of Hong Kong because of the increasingly popular pedagogy of task-based teaching and learning, students’ lack of communicative practice of form and personal form intervention in their short lessons, and the non-intimidating demand of recasts. Task-based teaching has become part of the official syllabi in Hong Kong primary schools since the mid 1990s and in secondary schools since 2001 (Carless, 2007). Nevertheless, Carless (2007), also Adams and Newton (2009) and Butler (2011), note that task-based teaching has not been incorporated well into lessons due to the large class sizes, exam-oriented assessment, and teachers’ unfamiliarity with employing it. Carless also argues in his study that task-based teaching in the Confucian-heritage culture of Hong Kong needs context-sensitive adaptation. For example, as Ellis (2009a) and Butler (2011) later suggest, there should be grammar focus to compensate the communicative focus of tasks, and to suit the traditional educational norm of grammar focus in Hong Kong schools. It is important to investigate if recasts, which are communication-appropriate to be embedded in tasks, are effective in bringing Hong Kong students’ focus to grammar during the tasks.

Furthermore, examining recasts in the context of Hong Kong is important to see if recasts facilitate students’ communicative use of grammar according to the task meaning. Yang and Lyster (2010) state that Chinese EFL learners are familiar with using discrete grammar form without a context because of their form-oriented education. Students’ limited class time (Yoshida, 2008) may also reduce their opportunities to experiment their use of form in communicative tasks. Limited class time may alternatively deprive teachers of giving personal feedback on students’
formal errors in communicative tasks. Teachers usually scaffold every meaning construction for students to save time (Philp & Tognini, 2009). The time-saving feature of recasts in giving direct reformulation to students may be a fine option for teachers to provide scaffold for learners’ use of form. Students may also benefit from the direct reformulation since it fits their habitual receptive EFL learning mode. It is therefore important to probe the effectiveness of recasts in the Hong Kong context.

Results and conclusions about the effectiveness of recasts have been diverse in the literature. L2 classroom studies (e.g. Lyster & Ranta, 1997; Panova & Lyster, 2002) observed that recasts were most frequently used by teachers because of their coherence with the interaction flow. The effectiveness of recasts in drawing students’ attention to the contrast between the error and target form is however less evident than other feedbacks which demand learners’ self-correction. As a result, recasts were often missed, or perceived as focusing on communicative meaning rather than form. The ESL student whom I taught before demonstrated his/her unawareness of recasts given within the ongoing interview between us. Experimental studies (e.g. Doughty & Varela, 1998; Han, 2002) on the other hand avoided the ambiguity of recasts by controlling the saliency of recasts. Some experimental studies showed the unbiased effectiveness of recasts, but some undermined the effectiveness of recasts in comparison to other feedback types due to their problematic control measures. Furthermore, recasts in and of themselves are both implicit and explicit, depending on their implementation. This may have been the main reason behind the controversy among studies yielding different results about recasts.

Although the different variables contributing to the specific results of recasts have been identified by past studies, there have not been sufficient studies comparing one type of recasts to another type of recasts along the same explicit-implicit
continuum. Drawing this kind of comparison may help control the variable of extreme
difference between the two feedbacks, which has not been widely controlled in past
studies. Moreover, studies on recasts have not conducted much investigation into the
effectiveness of recasts in the context of Confucian-heritage culture. The culture not
only entails limited exposure to the target language, but also receptive rote learning
relying on teachers’ predetermined syllabus input. The use of recasts in this context
could be a facilitative innovation, or a mission impossible due to the mismatch
between the communicativeness of recasts and learners’ form-oriented education. The
above uncertainty or interest motivated the present study to explore the use of
different types of recasts in a Hong Kong secondary school context.

The fundamental questions guiding the present study’s research proposal
were: 1) Are recasts feasible in the context of Hong Kong schools? 2) Are salient
recasts effective in facilitating Hong Kong learners’ learning? 3) How to balance both
the communicative and form focuses of recasts? There were three goals to aim at in
the present study. First, I wanted to examine the effectiveness of recasts in a
controlled experimental setting instead of a mass classroom setting in Hong Kong.
Second, I wanted to probe the relationship among recasts with different saliency
levels, learners’ learning, and learners’ perception towards recasts. Finally, I wanted
to implement recasts within a meaning-based task likely eliciting students’ use of past
tense, so as to balance the communicative advantage of recasts and the necessity for
grammar focus in a EFL context.

The present study chose the cognitive-interactionist second language
acquisition (SLA) approach as the theoretical framework. The decision is based on its
in-depth research on how language learning is derived from input (recasts) being
transferred to output (reformulation), through processing the linguistic knowledge in
the mind. Different methods can be used to facilitate the input processing involved, for example use of corrective feedback (e.g. Lyster & Ranta, 1997), input enhancement (e.g. Sharwood Smith, 1993), or processing instruction (e.g. VanPatten, 2004). Their effectiveness has been identified, tested and discussed in past SLA studies. The relation between the cognitive approach and foreign language learning in the Hong Kong context lies in the relative readiness of Hong Kong learners’ linguistic knowledge. Their linguistic knowledge of grammar rules has been more developed than their application of grammar rules in meaningful contexts, due to heavy doses of unidirectional grammar instruction from teachers in their EFL learning culture. The injection of recasts into the pedagogy may assist learners with mapping the required form with the corresponding meaning, which they are deficient in. Since the effectiveness of recasts has been said to depend on learners’ attention and perception and interlocutors’ implementation of recasts, the cognitive-interactionist theory guiding the implementation and hypothesized usefulness of recasts may advise the present study’s investigation of recasts.

The cognitive-interactionist approach however bears limitations in explaining how language is learned. The input-process-output model and learners’ attentional mechanism proposed in the theory is merely speculative, with no concrete evidence of their existence. Moreover, learners’ instantaneous noticing of recasts hypothesized by the theory is difficult to be ensured in both observable (verbal uptakes) and non-observable (no response) ways. Especially when the EFL Hong Kong students were not familiar with receiving personal feedback, recasts, and responding to recasts, the cognitive-interactionist approach, often employed in the ESL context, may turn out irrelevant in providing explanations. Still, this theoretical approach is adopted because it predominantly explains the way recasts may project effects. The suitability of the
theory for researching recasts in the Hong Kong context could be enhanced by the study's controlled methodology. The speculation involved in the theory and hence its applicability could be tested and reflected in the study's findings and analyses.

1.2 Thesis Organization

This thesis is composed of eight chapters. Chapter 1, the current chapter, introduces how the thesis is developed from a personal experience and interest in recasts. It has identified the significance of investigating the topic in the study’s context. From a brief report of the controversy in the field, the initial critical inquiry into the topic and justification for the theoretical framework selected have been stated. Some limitations of the theoretical framework are also noted.

Chapters 2 and 3 present a review of past literature on recasts. Different hypotheses and studies on the effectiveness of recasts have been discussed and critiqued. From the different studies on recasts, manifold variables, such as implementation of recasts, target form, learning measurement or test, learners' perception and level, task, modality and context, possibly affecting the resulted effectiveness of recasts have been pinpointed and examined. All the variables noted from the studies provide suggestions to the present study’s caution against any result distortion in its methodology. The review motivates the emergence of the study’s research questions.

Chapter 4 presents the research design of the present study. It delineates the quantitative and qualitative designs with reference to relevant past studies. The background of the target school and participants, the target form chosen, the task/test instrument design, the every step of conducting the data collection, and the
justification for the data analysis methods chosen are all included in the chapter. Ethical issues have also been mentioned. With the methodology specified, the study's research hypotheses are presented.

Chapters 5 and 6 analyse the quantitative and qualitative results of the present study respectively. Chapter 5 illustrates the use of different statistical methods and displays the results addressing the six research questions of the study. Beyond demonstrating the within-feedback group, between-feedback group, and across-test time quantitative results, findings about the patterns of verbs used by learners and variables within the feedback sessions are also shown. These are to provide a wide scope of quantitative analysis. Chapter 6 unveils the different issues raised by learners in their stimulated recalls of their perception of recasts and use of past tense during the feedback sessions. The issues are recurrent themes reported by learners across different feedback groups. Corresponding excerpts of stimulated recall interviews are presented.

Chapter 7 brings both the quantitative and qualitative data sets together for discussion. It summarizes the findings of each data set, and discusses their relation to the past literature and the present study's features. The two data sets are also compared and contrasted, by using one set to confirm and interrogate the other. Main issues which build the relevant discussions have been identified. The discussions involved aim at justifying the study's findings.

Finally, Chapter 8 provides a general summary and significance of the present study. The theoretical, methodological, pedagogical, and policy implications of the study are given in addition. The limitations of the research undertaken are also stated, which lead to suggestions for future research. Extension of current knowledge about
recasts by the present study is addressed. Reflections on the research process itself is made to bring the thesis full-circle and to a close.
Chapter 2 Literature Review I - Theoretical Variables of Recasts

2.1 Introduction

This chapter reviews existing literature about the different findings of recasts and rationales behind them. The necessity for corrective feedback has been argued differently by researchers. For instance, Long (1996, 2007) supports the need to give corrective feedback in interaction to facilitate L2 learning from negative evidence, exemplifying what is not acceptable in L2. Krashen (1982) however maintains that positive evidence exposure, exemplifying what is acceptable in L2, is already sufficient; and Truscott (1999, 2007) disapproves the facilitative effect of corrective feedback. Recently, Lyster and Saito (2010) affirm the significance and durability of corrective feedback in fostering L2 development through their meta-analysis of past studies.

Recasts have been observed the most frequently used in L2 classrooms among other corrective feedback types, as will be shown in the classroom studies discussed in this chapter. Besides this consistent finding of recasts across classrooms, other findings about recasts have been varying. From a psycholinguistic perspective, Long (2007) regards that the main potential advantage of recasts is their providing negative evidence to learners about their use of language in context. The common attentional focus shared by interlocutors is hypothesized to free learners’ attentional resources; form-function mapping may thereby be facilitated. With recasts being contingent upon learners’ own errors, the incorrect and correct utterances are juxtaposed for learners to make close comparison and contrast. These advantages are however not shared by the non-contingent characteristic of positive evidence.

Some other studies discuss other advantages of recasts, such as recasts being non-intrusive to students’ ongoing communication flow, brief in bringing the target
form forward, socially favourable in avoiding embarrassment in the case of students' failure in self-reformulation of errors, time-saving for EFL teachers to scaffold students’ use of form within their short lessons, fit for presenting models of lexical forms because of their positive evidence element in quickly retrieving students’ memory of the correct form learned, and fit for targeting forms which students have learned but have not automatized in online performance. Meanwhile, some studies illustrated recasts as ineffective because of their implicit implementation of embedding feedback in ongoing communication, and from which learners may not be able to perceive their correctiveness. Moreover, recasts have been critiqued as not engaging learners in pondering their own errors by straightforwardly presenting the correct models. These different views of recasts will be demonstrated and discussed throughout chapters 2 and 3.

Ellis and Sheen (2006) consider factors that may have caused the multifaceted identity of recasts in their re-examination of recasts. Factors such as saliency, intensiveness, repaired output, explicit or implicit feedback, positive or negative evidence, linguistic target, context/setting, learners’ orientation, and learners’ developmental readiness could affect the effectiveness of recasts on acquisition alone as well as in comparison with other feedback types.

Among all the aforementioned factors, Ellis and Sheen (2006) repeatedly conclude that saliency in overtly signalling the target form and intensiveness in consistently targeting a single form (e.g. Doughty & Varela, 1998; Mackey & Philp, 1998; Han, 2002) are fundamental to the effectiveness of recasts. For instance, how salient and intensive recasts are being delivered may guide learners’ orientation to the target form as the object or tool of language. Predictably, the more salient and intensive recasts are, instead of incidental and unplanned, the more likely learners
interpret recasts as explicit feedback driving metalinguistic awareness and negative evidence projecting the corrective purpose of recasts. Subsequently, the more explicit and corrective-like recasts are, the more likely learners notice and successfully uptake recasts. Similarly, the more explicit and corrective recasts are, the more form-oriented the context of learning appears. The more form-oriented the context is, the easier will learners notice any linguistic targets. With the help of explicit and corrective recasts giving both positive evidence to new forms and negative evidence to already-known forms, learners’ developmental readiness may comprehensively be catered. This catering for learners’ developmental readiness may however best happen with extensive and incidental recasts in classrooms, where no predetermined form is imposed on learners’ readiness.

Recasts with intensive, consistent and salient doses often occur in experimental settings. Doughty and Varela’s (1998) study implemented an explicit type of recasts to specifically target their science class learners’ use of simple past and past conditional. The kind of recasts given in their study is different from Long’s (1996, 2007) definition. Long defines recasts as given in one phase directly responding to learners’ preceding erroneous utterance—“a reformation of all or part of a learner’s immediately preceding utterance in which one or more nontarget-like (lexical, grammatical, etc.) items is/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 object” (Long, 2007, p.77). On the other hand, Doughty and Varela (1998) operated recasts, termed corrective recasts, in two phases—“via [1] repetition of the learner utterance with rising intonation [and stress on the error form]... plus [2] a recast [, when learner does not respond with using the target form,] providing the necessary target exemplar, either in contrast to a learner error or as a
model for a missing part morpheme, with recasts always delivered with a falling intonation curve [and stress on the target form]” (p.123).

Doughty and Varela’s recast operation can be interpreted as an extension of Long’s definition. Long implements recasts in a broader and more implicit way; and Doughty and Varela employ more restrictive (target only one feature—past time reference) and explicit (additional preceding attention-getting focus) recasts. Through comparing the treatment group which received focus-on-form instruction with explicit corrective recasts in addition to science content instruction and the control group which only received content instruction, Doughty and Varela found that corrective recast learners outperformed control group learners in the posttests. The results can be interpreted as deriving both from the explicitness, intensiveness and consistency of recasts and the reinforcing effect of the preceding repetition. Repetition has been defined by Lyster and Ranta (1997) as interlocutors’ repeating learners’ erroneous utterances verbatim; and classified by Lyster (2004) as one of the form-eliciting focus-on-form techniques called prompts. Prompts have been shown more effective than recasts because they force learners to generate their own reformulation (Lyster 2004; Ammar, 2008; Ammar & Spada, 2006). Therefore, it remains doubtful if the success of Doughty and Varela’s corrective recasts was attained exclusively through recasts or repetitions, or collectively through recasts and repetitions. All in all, Doughty and Varela’s complex operation of recasts, involving consistent frequency, intensive focus on one target, and saliency via preceding repetitions and intonational emphasis, made it difficult to isolate the intrinsic effect of each of these variables.

Mackey and Philp (1998) also conducted an experimental study investigating the effectiveness of recasts given intensively. In addition, they probed into whether learners’ developmental readiness and the presence of learners’ immediate response
to recasts affect the production and development of ESL question forms. The results showed that recasts had short-term positive effect on learners' use of question forms; and particularly advanced or ready learners performed better than their unready counterparts. Moreover, learners still benefited from recasts without immediate response, uptake in Lyster and Ranta's (1997) terminology, to recasts. This suggests that immediate response to recasts may not be an essential propelling learners' subsequent learning of the target form.

Han's (2002) small-scale experimental study overall found that individualized, intensive, and consistent recasts targeting learners' tense consistency (present or past tense) throughout their written and oral narratives generated facilitative effects on learners' production. From her finding of the effectiveness of recasts in keeping learners' tense consistency throughout their narratives, she identified four conditions for recasts to likely exhibit their effectiveness—individualized attention, consistent doses of recasts, intensive focus on form, and learners' developmental readiness for the form.

From the above three experimental studies, recasts can be implemented in certain ways by varying along the aforementioned variables, to deliberately maximize their facilitative effect on learners. Variables which affect the effectiveness of recasts are experimental setting, intensiveness in focusing on one target, consistent supply of recasts, individualized recasts, learners' developmental readiness for the target form, salient operation of recasts, and use of immediate response to recasts. These variables concluded from the above studies can be sorted into several macroscopic categories—context/setting, implementation, level or readiness, and measurement.

In view of these different factors shaping recasts, recasts could be made and found more effective than other feedback types. However, in their comparison to
prompts in past studies (e.g. Lyster, 2004), recasts were often found less effective because of not prompting learners’ responses. The following sections and the next chapter will scrutinise the different factors impairing or heightening the efficacy of recasts as advised by the different L2 theories and studies.

2.2 Input, Output and Interaction Hypotheses

Language learners start to learn a language in their first language (L1) acquisition. They acquire their L1 through continuous exposure to parental or caregivers’ input, for example recasting children’s erroneous or incomplete utterances (Bohannon et al., 1996; Farrar, 1990, 1992), during their meaningful communication. Krashen (1985) values the importance of input or positive evidence in facilitating learners’ language learning and establishes the input hypothesis. The input hypothesis theorises that second language (L2) learners may acquire their L2 through continuous exposure to the target exemplars presented as positive evidence in input. Only when input becomes comprehensible to learners, learners may then be ready to acquire the form that is next in line (i+1) for them to acquire. However, this relies heavily on learners’ autonomy to detect their own current stage as well as their next stage of knowledge in order to acquire the form concerned. Being sceptical of the presumed readiness of learners to attend to form at their own pace, Swain (1985, 2005) suggests a contrary proposal—output hypothesis. Swain’s output hypothesis promotes that learners need to be pushed to produce output and thereby move from semantic to syntactic and morphological processing. She argues that when learners only need to comprehend input, they may attend to the mere meaning and not be able to attain native-like accuracy. However, if they need to produce output, they may then attend to the form needed to be processed to encode the relevant meaning.
The failure of French immersion students in attaining L2 native-like accuracy has motivated Swain’s output hypothesis. Immersion students, studied by Lyster and colleagues (1997; 1998a, b; 2004), mainly focus on meaning and may only attend to form to the extent that is enough to communicate and be understood. To remedy their never-feel-needed focus on L2 form, Swain proposes that pushing learners to production may inspire them to notice the L2 form necessary for expressing the meaning, the gap between their interlanguage and the target form when interlocutors’ corrective feedback follows, and test their knowledge of form in their actual output. *Interlanguage* is learners’ developing state of language form before acquisition (Selinker, 1972). The push to attend to form in fulfilling the need of output is also implicated in Levelt’s (1989) speaking model. In Levelt’s model, the ultimate speech *articulation* stage is necessarily preceded by the *formulator* stage, where grammar encodes the preverbal meaning generated by the very first stage called *conceptualizer*.

Leeser (2008), a recent study testing Swain’s output hypothesis, found that the pushed output learners produced more attempts of past tense and less non-target-like forms. However, they had no significant performance concerning the target-like forms. Lesser relates this to the gradual and time-consuming acquisition of tense-aspect-meaning mapping (Bardovi-Harlig, 2000). This is precisely why Lesser used Doughty and Varela’s (1998) more realistic coding scheme of past tense attempts, instead of past tense accuracy, to closely capture learners’ interlanguage evidence (Doughty, 2004). Lesser also notes that a production task needs to be meaning-based for pushing learners to grammatically encode the meaning concerned. Lesser’s study may have yielded different results if learners’ pushed output were in the spoken mode. The effect of modality will be discussed in section 3.7 of chapter 3.
Although Swain’s output hypothesis seems to complement Krashen’s input hypothesis, both hypotheses somehow assume learners’ autonomy to process form out of meaning. This may be possible with learners who can notice on their own the gap between their interlanguage and the target language forms. However, for learners who cannot discover the gap alone, they may need help through negotiation of meaning during interaction with other sources, such as more proficient speakers or elaborated written texts. This is proposed by the interaction hypothesis (Long & Robinson, 1998).

Tasks with certain predictable negotiation of meaning occasions which provide modifications and simplification to enhance learners’ input comprehensibility without compromising their access to grammar forms may facilitate L2 form-function mapping. Long and Robinson further posits that negotiation also elicits negative feedback, such as recasts, to reformulate learners’ formal errors while retaining their meaning. Corrective feedback refers to “any reaction of the teacher which clearly transforms, disapprovingly refers to, or demands improvement of the learner utterance” (Chaudron, 1977, p. 31). Learners’ attention being drawn to forms may allow their noticing the mismatch between the different forms and hence cause focus on form. Such pedagogy may compensate learners’ deficient focus on form under the mere exposure to positive evidence.

Focus on form in Long’s (1996, 2007; Long & Robinson, 1998) definition concerns the allocation of attentional resources. This pedagogy suggests the benefit of occasionally shifting learners’ attention to forms by teachers or other learners under the circumstance of comprehension and production problems during interaction. The main focus is however said to be on meaning. Long distinguishes focus on form from focus on forms and focus on meaning; the former is midway between the latter two
which are the two extremes of systematic form-only focus and meaning-only focus respectively. Focus on form in Long’s term is therefore a kind of pedagogy of briefly implementing negative feedback to help allocate learners’ attention to form under the main backdrop of negotiation of meaning.

While Long approaches focus on form reactively by reacting to learners’ errors arisen in communication problems, other researchers have different approaches. Following Doughty and Williams’s (1998) early proposition of proactive focus on form, Ellis (2001) extends Long’s focus on form to include both planned and incidental form instructions. He generalizes both kinds of instructions as form-focused instruction (FFI). Planned FFI is a measure under which teachers or students predetermine a form to focus during communicative activities; and incidental FFI entails teachers’ focusing on forms arisen from the communicative need of a task. Both types of FFI are blended with communicative activities. On the other hand, Spada and Lightbown (2008) develop the concept of isolated and integrative FFI, and open an option of separating FFI from previous or subsequent communicative activities. Isolated FFI orients learners’ attention to the form intensively through structural syllabi, before providing opportunities for their use in communicative activities. Unlike Long’s reactive focus on form, integrative FFI can be proactive or reactive in drawing formal attention within communicative activities. It is also similar to Ellis’s (2001) planned and incidental FFI.

Both isolated and integrative FFI have been recognized by Spada and Lightbown as beneficial for different purposes. For instance, they synthesize past findings and identify that: 1) isolated FFI is conducive to develop foreign language learners’ precise and mature use of form in communicative tasks, due to their insufficient exposure to the target language outside classrooms, reinforced L1
influence from classmates sharing the same L1, as well as their habitual exposure to form-oriented classroom instructions; 2) integrative FFI is helpful for learners who are taught by the communicative approach, to develop fluency, accuracy and automatized use of the form outside classrooms where the target language prevails; 3) isolated FFI would provide enough saliency in targeting forms with easy rules but communicative redundancy in oral input, e.g. English third-person ‘s’; 4) integrative FFI may assist learners with complex forms which have less descriptive rules but more communicative value, e.g. English article; 5) unlike integrative FFI, isolated FFI may avoid interrupting learners’ ongoing communicative activities because of its separation from the flow; 6) isolated FFI may add dosage of formal input in meaning-based classrooms like immersion programmes.

To pedagogically address Dekeyser’s (2003, 2007) skill practice theory and Skehan (1998) and VanPatten’s (1996, 2004) support of information processing theory, Spada and Lightbown recommend the complementary use of both FFI types. Dekeyser upholds the use of practice to speed up the learning process of learners’ transferring their declarative knowledge to procedural knowledge. Declarative knowledge entails awareness of the rule explanation of form, and procedural knowledge is learners’ automatic use of the form to convey meaning and achieve certain communicative goals with or without the awareness of the rule behind (Anderson, 1983). This learning conception is established from the output point of view because of the goal of developing proceduralized knowledge. From the input point of view, VanPatten hypothesises that making form-and-meaning connections during real-time comprehension substantiates learners’ input processing and facilitates their learning. VanPatten defines form as lexical or linguistic items and meaning as content information. However, together with Skehan’s postulation of
human limited attentional capacities, VanPatten's (2004) *Primacy of Meaning* Principle argues that learners often prioritize meaning over form input comprehension when their attentional resources compete. Spada and Lightbown suggest that isolated FFI may help construct declarative knowledge and integrative FFI may automatize the use of the form to develop proceduralized knowledge. Moreover, the former can direct learners' attention to the form, make it cognitively manageable for learners to map form and meaning in instructional input, and later utilize it in output during communicative activities.

Focus on form techniques not only seem to bring learners formal input as well as attention to form via the external help of corrective feedback, but also opportunities to produce corrected output afterwards. The interaction hypothesis may therefore be comprehensive in bringing input, attention to form, and output together. However, as illustrated in the following descriptive and experimental L2 studies, there are issues concerning the effectiveness of different types of corrective feedback in different contexts/settings, different ways of measuring learners' learning, different characteristics of feedback and task, and different target forms.

This section has introduced the theoretical framework of interaction hypothesis and the pedagogy of focus on form derived from it. The learning element of interacting with interlocutors and written sources for comprehension and production assistance in Long's (1996, 2007) interaction hypothesis addresses the deficiency of not assisting learners sufficiently in Krashen's (1985) input hypothesis and Swain's (1985, 2005) output hypothesis. Long's focus on form is a reactive kind, other researchers, such as Doughty and Williams (1998), Ellis (2001), and Spada and Lightbown (2008) extend the practice of focus on form to proactively switching learners' attention to predetermined forms before they make any errors.
2.3 Corrective Feedback Classroom Studies

Lyster and Ranta’s (1997) content-based immersion classroom observation has been an early influential study devising an error treatment sequence model. It explains the interaction flow of their L2 teachers and young students. The model analyses the classroom interaction flow into: 1) from learners’ errors to either teachers’ feedback or topic continuation, 2) from teachers’ feedback to either learners’ uptake or topic continuation, and finally 3) from learners’ uptake to teachers’ additional feedback, reinforcement or topic continuation. Lyster and Ranta define uptake as “a student’s utterance that immediately follows the teacher’s feedback and that constitutes a reaction in some way to the teacher’s intention to draw attention to some aspect of the student’s initial utterance” (p.49); and repair as a student’s immediate successful uptake of the target form. Under the category of teachers’ feedback, Lyster and Ranta identify six different types: explicit correction, recasts, clarification requests, metalinguistic feedback, elicitation and repetition. Explicit correction straightforwardly tells learners which exact form is wrong; clarification requests are interlocutors’ request for meaning elaboration by usually asking “Pardon?”; metalinguistic feedback hints on the root problem with certain grammar rules; and elicitation guides learners to rethink their error form by pausing at the error for example. They further divide these feedbacks into two groups, teacher-generated and student-generated repairs, depending on the variables of uptake and repair. The difference of who generates the repair and whether uptake is encouraged are argued as contributing to the different rates of uptake following each of the above feedbacks.

Elicitation (100%), clarification requests (88%), metalinguistic feedback (86%) and repetition (78%) are termed negotiation of form and identified as prompting students to reformulate their errors on their own; whereas explicit
correction (50%) and recasts (31%) provide direct reformulation to students and students may or may not repeat. The difference in opportunities for students to produce uptake may have caused the above different uptake rates in Lyster and Ranta’s study; similar rates were found regarding repairs. The overall uptake rate in Lyster and Ranta’s classrooms was 55% and repair rate was 27%. Despite the least uptake rate, recasts were most frequently used. The authors emphasise that the distribution patterns of the feedbacks, uptakes and repairs only offer indications of learning instead of instances of learning.

Lyster later modifies his model by including one extra feedback, translation, and applies the model to a communicative language teaching setting with adult ESL learners (Panova & Lyster, 2002). Translation simply uses L1 grammar in L2. The communicative style of teaching in the study made language accuracy focus rather brief. However, Panova and Lyster predict that adult learners, who are more cognitively capable and motivated than children, would be more able to notice implicit and response-redundant recasts within communication. The issue of learners’ giving mere repetition-like responses following recasts, and the vague effectiveness of recasts in contributing to L2 development will be discussed after the discussion of L2 classroom studies.

Results of the study showed that recasts and translation were used most frequently. The uptake rate for each of the negotiation of form feedbacks accounted for more than 70% (elicitation 100%; clarification requests 100%; repetition 100%; metalinguistic feedback 71%); whereas recasts (40%), explicit correction (33%) and translation (21%) led to lesser uptake rates. Repair rates were also higher for feedbacks prompting the active participation of students to respond than those only triggering students’ repetition of the given reformulation. The authors admit that the
distribution percentages should be viewed with caution due to the disproportionally larger use of recasts than other feedbacks in both classroom studies (Lyster & Ranta, 1997; Panova & Lyster, 2002). The overall uptake rate in Lyster’s adult ESL study was 47% and repair rate was 16%.

Panova and Lyster (2002) make a point upon why recasts were used most frequently in their L2 classroom, even more frequent than Lyster and Ranta’s (1997) study. Although the learners were all adults, who are cognitively more capable than children in general, their beginning-level proficiency invited teachers’ decision to provide direct reformulation rather than challengingly prompting responses from students. Lyster and Ranta’s (1997) advanced class also tended to use less recasts than their other classes. The low proficiency of the adult ESL learners may have also caused Panova and Lyster’s lower overall uptake and repair rates than Lyster and Ranta’s, considering that adults were expected to be more attentive to feedback on form than children. This attribution of the appropriateness to use recasts in the ESL classrooms to learners’ low proficiency may also justify the use of recasts for EFL learners who have minimal exposure to the target language (Swan, 2005). However, the implicitness of the negative evidence in recasts may not be salient enough for less proficient and developmentally ready EFL learners.

Ellis, Basturkmen and Loewen (2001) also conducted an observational classroom study investigating focus on form techniques. Unlike Lyster and Ranta (1997), Ellis et al. (2001) recruited adult learners from different L1 backgrounds at an ESL private language school. This provides another perspective on how focus on form techniques are being employed in a L2 class with learners not sharing the same L1. The results showed that there were considerably high frequency percentages for both the students’ overall uptake (73.9%) and successful uptake (74.1%). Ellis and
colleagues concluded that the different L1 background factor, with the majority being Asian students, probably explained the much higher percentages than Lyster and Ranta’s study. Asian students are used to the grammar or instruction-oriented culture of learning; and learners at the private language school were believed to have higher motivation to learn because of their tuition fees-based education at the school. Furthermore, the first part of the lessons was grammar-oriented; this may have enhanced students’ anticipation of receiving form treatment in the following communicative part. Nevertheless, it has to be made clear that student-initiated focus on form episodes (FFEs) constituted the highest uptake rate; the next highest was uptake of responding FFEs; and the least uptake rate followed teacher-initiated FFEs. High successful uptake also occurred in the first two kinds of FFEs, and less than 50% in the third kind of FFEs.

The aforementioned different types of FFEs are another feature which distinguishes Ellis et al.’s (2001) study from Lyster and Ranta’s (1997). Ellis et al. additionally explore different focus on form techniques in which either students or teachers draw explicit attention to language features—student-initiated or teacher-initiated FFEs. Intentional learning operates when students ask questions or comment about confusing language features; or when teachers take time off from the ongoing interaction to fixate students’ attention to language features that they think they should learn. The highly active participation of students in student-initiated FFEs may have reinforced their following uptake or application of the form. Responding FFEs, which have been exclusively studied by Lyster and Ranta, are teachers’ incidental attempt to react to students’ previous erroneous utterances, via the primary use of recasts. Ellis et al.’s count of uptake is therefore broader in terms of including uptake not necessarily responding to a previous corrective feedback in student-initiated FFEs.
This broader count of uptake may not be comparable to that of Lyster and Ranta. Therefore, the highest student-initiated FFEs uptake rate, contributing to the much higher overall uptake rate in Ellis et al.'s study, may have created an illusion that their learners were more attentive to teachers' feedback. However, when specifically looking at Ellis et al.'s responding FFEs, the uptake (75.3%) and successful uptake (78.6%) percentages were still higher than Lyster and Ranta's uptake (55%) and repair (27%) rates. The most frequently used responding FFEs feedback, recasts, also triggered higher uptake (71.6%) and successful uptake (76.3%) rates than Lyster and Ranta's study. All in all, the higher uptake and successful uptake rates of recasts in Ellis et al.'s study may notably be related to its more form-oriented context.

Besides taking a different context and wider scope of examining different FFEs, Ellis et al. also extend Lyster and Ranta's study by investigating the influence of the different characteristics of FFEs on uptake. Concerning the source of the FFEs, form-oriented episodes appeared more frequently than their meaning-oriented counterparts. However, the uptake rate was higher for meaning-oriented FFEs where lexical form dominated the focus. The source of FFEs had no significant effect on the successful uptake rate. Simple FFEs involving single exchange were the majority, but complex FFEs involving multiple exchanges significantly helped learners' uptake and successful uptake. The directness of the FFEs, use of implicit recasts in responding FFEs or explicit provision of form in student-initiated FFEs, did not differ significantly in influencing uptake and successful uptake. Both implicit recasts and explicit focus led to high uptake and successful uptake rates. Regarding linguistic focuses, grammar and lexis were dominant. Uptake was not affected by which form
was being focused; but successful uptake tended to be resulted from pronunciation focus.

From the above discussion of the three L2 classroom studies, the variable of context has been recurring. Firstly, Lyster and Ranta (1997) studied young L2 learners in a content-based immersion program; the dominant focus on content meaning and children’s unreadiness may have engendered the low uptake and repair rates. Secondly, Panova and Lyster (2002) refined and applied Lyster and Ranta’s earlier corrective feedback interaction model to an adult ESL communicative classroom setting. The not much higher rates of uptake and successful uptake than the earlier study was speculated to be due to the adult learners’ beginning-level proficiency, though adults are supposed to be more cognitively ready than children. Thirdly, Ellis et al.’s (2001) more extensive study, which explored the different focus-on-form techniques and their different characteristics in effecting uptake and successful uptake, related the higher uptake and successful uptake rates to their more motivated adult ESL learners and form-oriented teaching. Ellis et al.’s findings inform the need to further investigate the context and focus-on-form characteristic variables in affecting learners’ uptake and successful uptake.

Loewen (2004) used a similar population from a different L2 private language school and examined which particular variables induced a causal relationship with learners’ uptake and successful uptake. These were argued to consolidate Ellis et al.’s (2001) claim that context is a significant variable and complement their observation of the connection between focus-on-form characteristics and the resulting uptake. Loewen states that his study attempts to fill the gap of studies which only probe into the increasing likelihood of uptake triggered by certain focus-on-form characteristics. His study, also largely with Asian adult ESL students in the form-oriented school,
ventures to find which specific characteristics causally drive learners’ production of uptake.

Loewen (2004) obtained the overall uptake rate of 73% of all the FFEs, nearly equivalent to that of Ellis et al. (2001), and 66.1% successful uptake. Eight FFE characteristics, type, linguistic focus, source, complexity, directness, emphasis, response, and timing, were analysed to draw the degree of their causal relationship with learners’ uptake and successful uptake. Apart from sharing the first five characteristics with Ellis et al., Loewen also examined emphasis (combination of complexity and directness), response (provide reformulation or elicit reformation from students), and timing (immediate or deferred response). Among these characteristics, complexity, timing and response were found significantly predicting uptake. Specifically, FFEs with more than one response move and immediate response elicitation were causes of uptake; and FFEs with heavy emphasis (direct and/or complex), formal feature source, and reactive feedback type (most frequent) led to successful uptake in addition.

The facilitative variables of motivated adults and form-oriented context have been affirmed in Loewen’s (2004) study, which gained comparably high percentages of uptake and successful uptake with Ellis et al.’s (2001). Moreover, with the expansion of studying the causal relationship between different FFE characteristics and learners’ uptake, Loewen’s research inspires that different variables can possibly exert direct effect on focus on form. Notwithstanding the contribution he has made to the causal effect of the different FFE variables, Loewen admits that the effectiveness of uptake in impacting subsequent L2 learning still remains questionable.

Loewen’s (2005) following study attempts to clear the doubt of the learning effect of uptake. Using the same setting, he implemented a non-descriptive approach
to test whether learners could produce the same form in uptake in the two subsequent posttests. The posttests were designed to prompt learners’ recall of the same individual items that had been produced in their previous uptake of FFEs, using the same FFE meaning context. The use of individualized tests and same items was to secure occasions for learners’ later production of the form. The results illustrated that learners were able to recall the linguistic items in previous FFE uptake almost 50% correct in the 1-day-after posttest and nearly 40% correct in the 2-week-after posttest. Significantly, more correct responses in the tests emerged for items being successfully taken up in the FFEs, indicating the key of the quality of uptake. These may suggest that uptake, or largely successful uptake, is an adequately accurate measure of the effectiveness of incidental focus on form, in predicting learners’ later use of the form.

However, Loewen (2005) recognises the spontaneity limitation of the tests. The tests only required learners to recall the exact individual linguistic items produced by the respective learners in their previous uptake of FFEs. Recalling the same linguistic items under the same meaning context may in fact be testing learners’ memory instead of their use of the form. Addressing this problem, he calls for future research to use tests prompting learners’ more spontaneous use of the relevant linguistic form. This may help investigate learners’ subsequent learning of FFEs beyond the level of immediate uptake and recall of discrete items. Moreover, to examine the possible effect of learners’ noticing without producing uptake, Loewen suggests that comparing learners’ noticing with and without uptake production through more controlled experimental and stimulated recall studies would be an avenue for future research.

Although all the classroom studies above demonstrated that learners responded to the various types of corrective feedback, they acknowledge that immediate uptake
may not be sufficient to serve as evidence of learning. The reason is attributed to the
different quality of uptake and opportunities for uptake following different types of
feedback. However, uptake is not totally futile. The aforementioned L2 classroom
researchers consider uptake as indicative of learners' instantaneous noticing of certain
form-and-meaning connection, with Loewen (2005) finding additionally that uptake
predicts subsequent learning from previous FFEs. Swain's (1985, 2005) pushed
output theory justifies that immediate uptake shows learners' noticing the mismatch
between their errors and the target form. This may clearly be the case for feedbacks
eliciting students' responses. For feedbacks providing direct reformulation like
recasts, arguments have been diverse.

First of all, Mackey, Gass and McDonough (2000) found that recasts with
learners' following repetition were more probably perceived as corrective and lack of
response might represent the ineffectiveness of recasts as feedback. Despite such
finding, feedbacks targeting morphosyntactic forms were least perceived as feedbacks
in learners' stimulated recalls; and most morphosyntactic feedbacks were recasts.
More radically, Gass (2003) contends that uptake may only be learners' "mimicking"
(p.236) exemplars without noticing the negative evidence in recasts; and may not
show their "true understanding" (p.236) of teachers' reformulations.

Pinker considers that negative evidence should exist, be usable, used and
criteria for negative evidence to play a role in L1 acquisition to L2 acquisition. He
specifies that whether recasts are noticed and perceived by learners as negative
evidence determines the effectiveness of recasts in drawing the mismatch between
learners' errors and the target forms. This corresponds to Pinker's criterion of
usefulness. Learners' incorporation of the corrected form into their utterances is
identified by Long as Pinker's criterion of using the negative evidence. However, as shown in the above L2 classroom studies, recasts often triggered the least learners' incorporation or uptake rates of the correction; and even when learners did show uptake, as Gass (2003) argues, it may only be a mechanical repetition not involving noticing. These phenomena seem to make recasts run against the usefulness and utility criteria for their negative evidence to exhibit influence. The feature of providing direct reformulation and not pushing self-correction may have made learners think that it is redundant to give a response.

The inappropriateness to give a response after recasts within communicative interactions may have been another cause of absence of response. Based on Pinker's (1989) criteria of the role of negative evidence, Braidi (2002) examined the existence and use of recasts during interactions between native and non-native speakers. Recasts were found to exist in their interactions, particularly more in extended than one-signal negotiations and more in response to multiple-error than single-error utterances. Braidi speculates that learners may have needed recasts most in difficult situations of extended negotiations and multiple-error utterances. Concerning utility of recasts, more learners were found incorporating the correction when only those interaction scenarios which were considered appropriate and possible for learners to respond were counted. Braidi's findings therefore confirm the existence criterion for recasts to play a role in L2 learning, as well as introduce appropriateness or possibility to respond to recasts during ongoing interactions as a factor driving learners' incorporation of the form recast.

This section has discussed the effectiveness of recasts in L2 classrooms and the possible reasons behind the different findings. It first of all delineates Lyster and Ranta's (1997) immersion classroom observation, then Panova and Lyster's (2002)
ESL adult classroom study, and Ellis et al. (2001) and Loewen’s (2004, 2005) ESL adult private school study. The different uptake and repair rates observed from the different classroom studies can be attributed to the different levels of students and the different focuses of classrooms. Compared to Lyster and Ranta’s early study, Panova and Lyster’s beginning-level adult learners yielded lower rates of uptake and repair than Lyster and Ranta’s young learners. Ellis et al. and Loewen found higher uptake and repair rates than Lyster’s studies, because of their form-oriented contexts. Loewen’s two studies further probe the relationship among interaction variables, uptake and subsequent learning. Although uptakes may indicate learners’ learning, they may be learners’ empty repetitions. Furthermore, uptakes may not be appropriate responses in ongoing communication. Therefore, uptakes may not be an accurate measurement of learning.

2.4 Noticing

From Pinker (1989) and Long’s (1996) stance on the usefulness and utility criteria of negative evidence, the potential relationship between learners’ noticing, perception, and response emerges. However, the relationship may be a loose one when learners’ response does not carry any noticing of the negative feedback. Learners’ perception may serve to insightfully reveal whether their response to feedback reflects noticing.

Yoshida (2010) employed stimulated recalls to probe teachers’ perception of students’ responses to the feedback given and students’ perception of the feedback. In occasions when teachers perceived that students noticed the feedback via their acknowledgement or successful uptake, students recalled that they in fact did not notice the feedback. Mismatches between teachers’ and students’ perception
happened: 1) when students responded to teachers’ correction but did not understand the difference between their errors and the target form; 2) when teachers thought students’ responses indicated their noticing of the feedback, but the nature of students’ problem had not been targeted, leading to students’ zero noticing. Students reported that their responses were given to avoid embarrassment in breaking the ongoing interaction flow with the teachers. They were actually thinking about their own errors or how to construct sentences. Moreover, teachers revealed that they usually gave implicit feedback to avoid causing strains to students. Yoshida’s data may warn researchers that learners’ responses of either acknowledging or successfully taking up the feedback could come from other purposes than noticing. Learners’ perception may help unveil their covert interpretation of the feedback, but there are issues about accessing learners’ perception, which will be raised in section 3.2 of chapter 3. Learning, instead of verbal uptake, is perhaps a more reliable and substantial mirror of learners’ noticing.

Even when learners notice the contrast between their errors and the target form that teachers’ recasts bring forth, their preference for other feedback types may mediate the learning effect that recasts can give to learners (Yoshida, 2008). Yoshida’s (2008) previous study used stimulated recalls and unveiled that the learners actually preferred feedback (e.g. elicitation) which allows them more time to think over the error-target contrast and sense of achievement in producing self-correction. However, the teachers revealed that they needed to use recasts despite their parallel preference for feedback (e.g. elicitation, metalinguistic feedback) guiding learners to self-correct and extending explanations for the form. Because of the lesson time constraint and cognitive style of the learners, the teachers had to stick to recasts which directly provide the correct form and thereby reduce the learners’ embarrassment.
when not knowing how to self-correct. It can be inferred that recasts are pedagogically and socially suitable to classrooms which have tight lesson time and learners whose cognitive capacity may not permit them to initiate self-corrections. In inevitable consideration of these pedagogical and social factors, learners’ learning may as a result be compromised.

Noticing has been regarded as transforming input to input attended, called *intake*, for further processing (Corder, 1967; VanPatten, 1996, 2004). The position of noticing in second language learning has been theoretically sound but empirically less feasible to test. Schmidt (1990, 2001) proposes the *noticing hypothesis* to state that awareness is necessary to L2 learning. Robinson (2001a) concurs with Schmidt in terms of the essentialness of deploying attentional resources; he also adds that active rehearsal in the short-term memory facilitates further processing. However, Tomlin and Villa (1994) and Gass (1997) argue that initial registration of input or *detection* is sufficient; and unawareness can lead to learning. It has to be noted that the form that these researchers deal with is already learnt. Schmidt stresses that awareness is crucial when learning new L2 form. He classifies *awareness* into different levels: 1) *noticing*, the subjective account of the surface L2 form; and 2) *understanding*, a deeper form of awareness which concerns the learning of the abstract L2 rules for generalization.

Connecting noticing with feedback, Mackey (2006) explored the relationship between interaction feedback (negotiation and recasts), noticing, and L2 development via a small-scale quasi-experimental study. It triangulated multiple introspective measures to probe into instances of learners’ noticing classroom feedback, and used a pretest-posttest design to gauge the effectiveness of noticing on their subsequent learning. Triangulation was implemented to counterbalance the different extraneous variables affecting the validity of different noticing measures. For example, reflective
journal and questionnaire could take a long time to complete, think-aloud protocol may trigger inadequate responses due to online time pressure, and stimulated recall may guide learners to create new thoughts instead of recalling past events. Although the study tried to minimize the threat to validity, Mackey (2006) notes that only occurrences of noticing feedback and their relationship with learning can be drawn. The substantial causation between noticing and learning cannot be easily observed, because no reported noticing of the form in feedback does not necessarily equate lack of noticing and hence zero learning.

The ESL forms targeted by the study’s interaction activities between teachers and learners were question forms, plural forms, and past tense. Learners’ development was measured through their production of the forms elicited by the pretest and posttest tasks. Results concerning learners’ noticing and subsequent learning were selective according to the different forms. Interaction feedback was found to promote learners’ reported noticing of question forms for the most percentage, the percentage for plural forms came next and past tense the least. The number of learners who reported noticing feedback to question forms and developed was the highest among the three forms. Three possible reasons were stated to explain these findings: 1) the syntactic movement in question forms appears more salient than past tense involving morphological change; 2) question forms carry more communicative value than past tense; and 3) question forms were more often negotiated and the two other forms were more often recast. Mackey cautions that the multiple introspective measures and more question forms being negotiated than recast may have in fact heightened learners’ noticing and learning by frequency and pushed output.

Apart from external influences of context, task design and target form, internal cognitive constructs may also affect learners’ noticing and benefiting from recasts.
Trofimovich, Ammar and Gatbonton (2007) investigated the impact of phonological memory, working memory, attention control, and analytical ability on ESL French learners’ noticing of correct possessive determiners (morphosyntactic) and intransitive verbs (lexical) in recasts. Learning was measured in the subsequent post-test and delayed post-test. Firstly, phonological memory refers to a learner’s ability to temporarily store spoken utterances in the short-term memory or a working memory subcomponent called phonological loop (Baddeley & Logie, 1999). Being capable to encode and retain the spoken utterances in the phonological loop may predict learners’ ability to repeat and learn the form in recasts. Secondly, working memory defines learners’ ability to concurrently process and retain the verbal form needed for a task. It may in other words designate the amount of verbal information learners can attend to simultaneously and the scope of recasts that can be noticed, processed and analysed. Thirdly, analytical ability reflects learners’ grammar sensitivity. This may allow learners to focus on the form in their speech as well as the addressors’ speech in the form of recasts. Last but not least, attention control denotes learners’ ability to efficiently allocate attention resources to different linguistic aspects or cognitive processing tasks. A relevant scenario would be how capable learners distribute attention to perceiving recasts and subsequently encoding their own meaning.

The study employed contrived recasts through prior recording instead of spontaneous recasts in face-to-face interaction. This may have attenuated the reactive nature of recasts. The process went as learners answered pre-recorded simple questions eliciting their one sentence with the target form describing the picture on the computer screen. Pre-recorded recasts were given to both correct and incorrect sentences. Non-corrective recasts to the correct answers may inevitably enhance learners’ noticing the form. Afterwards, yes/no questions were asked to examine if
learners noticed the difference between their form and the target form brought by previous recasts. Following were an immediate post-test and a few minute delayed post-test prompting the same answer describing the same picture. Tests measuring the aforementioned four cognitive processing constructs were administered to see their correlation with noticing and learning of recasts.

Results showed that learners were more likely to notice their lexical than morphosyntactic errors. Moreover, their production accuracy in the two post-tests was greater than that of the pre-test. Factors of proficiency, attention control, analytical ability, and phonological memory were found predictive of learners' accuracy but not noticing. The insensitivity of these factors to influencing learners' noticing was attributed to the saliency of the recasts and tasks given. Recasts were systematically given by the recorder and the task simplistically involved the same one-sentence answers to the same pictures throughout. The task was also non-face-to-face interactive, which avoided unpredictable demands on learners imposed by interlocutors' unclear meaning. Learners may not have needed much cognitive processing to rely on. Among the effects of the three cognitive factors, both analytical ability and phonological memory exerted more effects in the delayed post-test when recasts no longer existed to sensitize learners' attention to the form and orally deliver the form. At that moment learners would really need to depend more on their cognitive efforts. This may precisely be why immediate response to recasts often did not reflect learners' cognitive processing. The effect of attention control was indicated by learners' attention distributed to both the morphosyntactic and lexical forms, as shown in their production accuracy.

This section has looked into the internal factor of noticing in prompting learners to learn the form and reflecting learning through learners' responses.
Different ways of revealing learners’ noticing have been shown. Yoshida (2008, 2010) used stimulated recalls to inform students’ unawareness of the feedbacks given and their preference of self-reformulation. The teachers and students also recalled choosing recasts because of their non-intrusiveness to the ongoing communication and time-saving feature of giving reformulation directly. Mackey (2006) also used stimulated recalls to probe the effectiveness of negotiation and recasts, but with the triangulation of other introspective measures. Her investigation into learners’ noticing unveiled that their degree of noticing differed according to the different target forms. Trofimovich et al. (2007) meticulously tested the effect of different cognitive factors on learners’ learning.

2.5 Response to Recasts

The variable of pushed output (Swain 1985, 2005) after feedback has been conceptualized as securing opportunity for learners’ response and driving their noticing. As informed by the aforementioned classroom studies, immediate response to recasts may not be required by the ongoing discourse and may not be qualitatively beneficial to learners even if repeated. Rather than absolutely denying the value of immediate response to recasts, Egi (2010) argues that its different qualities may contribute to learners’ noticing of the corrective intent of recasts. The study made use of stimulated recall to uncover learners’ perception of their recast episodes with uptake, repair and modified output of the errors. Through probing learners’ inner thoughts, Egi assumes that the value of immediate response to recasts can be noted to a certain extent. Different from Lyster and Ranta’s (1997) identification of uptake and repair, Egi puts forward the new category of modified output. She defines it as either learners’ targetlike or non-targetlike modification.
The study’s results illustrated that learners perceived recasts as corrective in their report more when they produced uptake in the feedback episodes than those without uptake. Moreover, learners were not only more likely to report the corrective intent of recasts but also the gap between their errors and the target forms when they had repaired their utterances. Learners also explicitly recalled recasts as corrective as well as identified the difference between their errors and the target forms more when they modified their utterances. It is noteworthy that learners could recall both the intent and function of recasts when they modified their utterances targetlike as well as non-targetlike. This may consolidate Swain’s (1985, 2005) output hypothesis that any output, whether it is targetlike or needs further testing and feedback, is facilitative to learners’ cognitive processing.

Egi’s (2010) findings unveiled the different levels of cognitive activity that learners engaged in when exhibiting the different qualities of immediate response to recasts. Among the different qualities of response, repair and modified output, which fostered learners’ noticing the contrastive gap between target and non-target forms, would be particularly useful in converting input to intake (Saxton, 1997). Investigating the different qualities of response to recasts may move the field forward by defending the usefulness of immediate response to recasts. Egi envisions that one holistic study with verbal report, immediate performance and subsequent acquisition data would comprehensively trace learners’ L2 development in relation to their immediate response to recasts.

Conventionally, response to recasts has been understood as learners’ immediate verbatim repetition of the target reformulation. McDonough and Mackey (2006) state that response to recasts can on the other hand appear as primed production. They specifically investigate the relationship between recasts, EFL
learners’ responses to recasts, and development of English question structures. 

Repetition was defined as learners’ use of the question structure verbatim as that of interlocutors immediately after recasts; whereas primed production was learners’ use of the question structure in different wording either in immediate turns or several turns after recasts. Immediate primed production was in fact less frequent than delayed primed production in the study. Recasts and learners’ different responses were elicited through conversational information gap and exchange activities. A pretest and a posttest were designed in the form of oral production to gauge learners’ question development. Results illustrated that both recasts and primed production were significantly correlated with L2 question development; but repetition was not. The study therefore suggests that delayed responses to recasts rather than immediate ones relate to L2 development. The superior effect of incorporating recast in extended production over repeating recasts in facilitating subsequent learning was later confirmed by Nassaji’s (2011) study.

Deriving from the above findings about syntax, target production of grammar forms in different lexical or morphological constructions after recasts may also demonstrate learners’ learning. This would be through learners’ applying the abstract rule of the form to other instances. N. Ellis’s (2005) proposal of associative learning which turns explicit knowledge to implicit knowledge may offer relevant elaboration. This will be discussed in section 3.3 of chapter 3. This indication of learners’ learning needs to be considered in L2 research as well, apart from taking learners’ exact repetition or recall of the form as indicating or measuring learning (Ellis et al, 2001; Loewen, 2004, 2005).

McDonough and Mackey (2008) further investigate learners’ use of the same or different lexical or morphological forms in constructing their primed production of
the target structure. The departure from their previous study (McDonough & Mackey, 2006) is that this later study particularly excluded the possibly intervening effect of recasts and looked into the intrinsic effect of learners' primed production. The finding showed that learners' primed production had significant effect on their subsequent learning of English question formation. Apart from this crucial finding, the authors also discovered that learners' primed production of the target form in different lexical and morphological constructions led to learning more significantly than their primed production verbatim as interlocutors' model priming. This supports Panova and Lyster (2002) and Gass's (2003) claim that mere repetition of the form contributes no significant learning. All in all, the effectiveness of applying the target form in various lexical or morphological exemplars on signaling and facilitating learning has been consolidated in the study.

The skill acquisition approach argues that the extensive application of the abstract rule of the form or strengthened development of the abstract rule representation provides evidence of proceduralized skill or skill automatization (DeKeyser, 2001, 2007). Learners' repetition is on the other hand considered practising the knowledge for its later proceduralization (DeKeyser, 2007; Nassaji, 2011). This may lend some support to the learning value of learners' repetition of recasts. The skill approach also explains the effectiveness of prompts (DeKeyser, 1998, 2001, 2007). To facilitate both communicative and controlled practices, prompts may help through scaffolding opportunities for controlled practice in communicative activities by requiring the use of specific forms (Lyster, 2007). Pushing output in prompts may assist with learners' transition from declarative to procedural knowledge (Ranta & Lyster, 2007).
This section has introduced a different kind of learners’ response to recasts which can be used to measure learners’ learning from recasts. Other than immediate response to recasts, which Egi’s (2010) study showed that the different qualities of learners’ immediate response may contribute to learners’ noticing of recasts, incorporating recasts in extended production may also reflect learners’ learning from recasts. McDonough and Mackey (2006, 2008) and Nassaji’s (2011) studies have demonstrated that primed production, which is learners’ reformulation in several turns after recasts, may reflect learners’ subsequent learning better than immediate response.

2.6 Prompts and Recasts

Besides the concern of facilitating learners’ noticing, another limitation of recasts is their absent elicitation of learners’ repair. Such limitation of recasts often reflects the outstanding corrective potential of prompts. Prompts have been found more effective than recasts in eliciting students’ active engagement to respond and fostering L2 learning in different classroom studies (Ammar & Spada, 2006; Ammar, 2008; Ellis et al., 2006; Ellis, 2007; Lyster, 2004; Yang & Lyster, 2010). Lyster and Saito (2010) have thoroughly examined the effectiveness of corrective feedback in 15 L2 classroom studies in their meta-analysis; and concluded that prompts have been found carrying larger effect than recasts across different years. They therefore suggest that prompts may be a reliably better pedagogical measure in tackling learners’ interlanguage. They attribute this difference in effect to prompts’ eliciting or cluing feature more than their explicitness in bringing negative evidence forward. They further explain that prompts, with their response-eliciting or cueing feature, would be more pedagogically sound because teachers could in fact overtly apply cueing
techniques during their classroom interaction with students. However, applying explicitness in delivering feedback may not be necessarily conspicuous to students depending on their perception (Spada & Lightbown, 2008). This section will analyse theoretically and empirically the difference between prompts and recasts.

Lyster (2004) advances from his early descriptive studies on communicative contexts to a more controlled approach, quasi-experimental, to examine the effectiveness of recasts in comparison to prompts in *form-focused instruction (FFI)* classrooms. FFI combined with prompts were found to be more effective than combined with recasts in written tasks, and in oral tasks to a smaller degree. According to Ellis (2001), FFI attempts to either draw learners’ attention to language features implicitly and incidentally during ongoing classroom communication, or explicitly and structurally based on education syllabi. Lyster’s (2004) study was motivated by his French immersion students’ grammatical inaccuracy in achieving native-like competence, predictably due to their program’s content-based or meaning-oriented approach. To examine whether doses of formal focus can boost immersion students’ French grammatical gender accuracy, Lyster (2004) applied both the planned element of FFI and the incidental supplement of oral corrective feedback—prompts and recasts.

Recasts are feedback targeting learners’ formal errors arisen in their message without demanding them to reformulate the errors by themselves; whereas other feedback types, elicitation, metalinguistic feedback, repetitions, and clarification requests, prompt learners’ self-correction without offering them a direct solution as recasts (Lyster & Ranta, 1997; Panova & Lyster, 2002). Prompts, such as elicitation, metalinguistic feedback, repetitions, and clarification requests, have a corrective rationale which can be explained by Swain’s (1985) output hypothesis. The
hypothesis posits that the move of pushing learners to produce the form targeted in feedback can assist them with noticing and revising or testing their non-target formal hypotheses (Swain & Lapkin, 1995). From prompts’ hands-on opportunity or practice for learners to produce a modified form, Lyster (2004) concludes that the advantage he found of prompts over recasts in ensuring learners’ output may help realise learners’ proceduralization (Anderson, 1983; Johnson, 1996; Johnson & Jackson, 2006), and thereby acquisition of French grammatical gender.

According to the skill acquisition approach (Anderson, 1983), there is a difference between declarative knowledge and procedural knowledge. The former is “knowledge about” (Johnson & Jackson, 2006, p.534), often represented by the rule-based system (Skehan, 1998); and the latter is “knowledge how to” (Johnson & Jackson, 2006, p.534), often represented by the exemplar-based system (Skehan, 1998). Lyster (2004) argues that prompts, by mandating and thereby training learners to output the target form hinted, can encourage learners to retrieve the already-known form in rule-based representation from long-term memory and apply it in its correct production. Learners may then proceduralize their declarative knowledge for later automatized usage of the form.

Though recasts do not force learners to reformulate output to practise proceduralizing the form in focus, recasts being employed with frequency and perceptual salience may enable learners to notice (Schmidt, 1990, 1993) the target models or positive evidence (Leeman, 2003) exemplifying the proceduralized use of the form. For illustrating the effectiveness of recasts with frequency, there is the success of intensive recasts in Mackey and Philp (1998) and Han’s (2002) studies discussed above, which focus on one form consistently in an experimentally controlled setting. For demonstrating the effectiveness of recasts with salience, there
are Doughty and Varela's (1998) corrective recasts with tonal change to the error form in a preceding repetition prompt, to help learners identify the “locus of the error” (Ammar & Spada, 2006, p.563) before any necessary recasting.

Some may argue that it is actually the frequency or perceptual salience variable favouring the effectiveness of certain feedback. The frequency variable which enhances learners’ noticing the form in focus by increasing the amount of exposure (Schmidt, 1990, 1993) may have caused the effectiveness of prompts over recasts in Lyster’s (2004) study. The form-focused instruction (FFI) session provided to all the three experimental groups (FFI and prompts, FFI and recasts, FFI only), prior to the implementation of prompts and recasts for these two groups, was designed to include: 1) typographically enhanced text noticing activities, 2) inductive rule-discovery and metalinguistic explanation awareness activities, and 3) analytic and fluency-based practice activities. The effectiveness of FFI versus none has been confirmed in the study with all the three FFI groups outperforming the control group. The FFI employed in the study, especially the inductive rule-discovery and metalinguistic explanation awareness activities, resembles one of prompts’ purposes of giving metalinguistic clues to guide learners’ self-correction. Therefore, prompts’ benefit over recasts may have been because of such frequent prompting with the prompting group having received both prompt-like FFI and feedback in the study.

The frequency effect of prompts on the prompted group in Ammar and Spada (2006) and Ammar’s (2008) studies investigating the comparative effect of prompts and recasts may have also led to their false superiority over recasts. The studies’ FFI session prior to the prompting and recasting treatments given to the two experimental groups was clearly prompt-like as shown in the guidelines provided to the participating teachers. For instance, the rule of thumb in the form of question
suggested for teachers to teach in the FFI session was “Whose _____ is it?” or “Who does it belong to?” (Ammar & Spada, 2006, p. 572); and two of the examples given to the prompting teacher to follow were “No. Whose hat is it?” (Ammar & Spada, 2006, p. 572) and “Who does the hat belong to?” (p.573). Though the studies’ result showed that prompts were more effective than recasts in facilitating low proficiency learners’ use of the target form, it may be difficult to see the exclusive effect of prompts over recasts for low proficiency learners because of the frequency variable favouring prompts.

Ammar (2008), part of the same larger study as Ammar and Spada (2006), concentrates on analyzing the comparative speed of prompted and recast learners in retrieving the third person possessive determiners in English. Ammar (2008) measured the extent to which prompts and recasts automatize learners’ retrieval of the target form by their speed of reacting to the computerized fill-in-the-blank task. The computerized task asked learners to choose an answer out of three options to fill in the blank of a sentence given each time. The reaction time of learners’ choosing an answer and turning to the subsequent screen was recorded. This was to analyse their speed of retrieving the form previously targeted by feedback, and the different feedback types’ extent of automatizing learners’ retrieval of the form.

The result showed that prompts sped up or automatized learners’ retrieval of the form more than recasts did. This seemingly demonstrated prompts’ superiority in automatizing or proceduralizing learners’ form retrieval. However, the variable of frequency may have distorted the results again. The fill-in-the-blank questions, for example “The boy is playing with _____ sister” (p.192), in the computerized task carried the same elicitation quality as prompts. The reoccurrence of prompts eliciting the form in the task may have enhanced prompted learners’ noticing (Schmidt, 1990,
1993) and expecting the use of the form targeted by prompts before. Therefore, it could have been the frequency variable biasing the effectiveness of prompts over recasts in speeding up or automatizing learners’ retrieval of the form.

Moreover, the use of such controlled and explicit context for gauging learners’ implicit automatized retrieval of the target form may have undermined the validity of the finding. As exemplified above, the fill-in-the-blank task limited learners’ focus and production to the retrieval of possessive determiner by presetting the sentence context, without letting learners construct the meaning. Such biased orientation to the form may have called for learners’ obligatory “explicit memory-based performance” (Norris & Ortega, 2000, p.483) or reliance on rules, rather than implicit automatized use of the form through building their own meaning. In their meta-analysis of past studies on the effects of explicit and implicit L2 instructions, Norris and Ortega (2000) raise the following critique:

...the measurement of change induced by instruction is typically carried out on instruments that seem to favour more explicit types of treatments by calling on explicit memory-based performance...over 90% of the dependent variables required the application of L2 rules in highly focused and discrete ways, while only around 10% of the dependent variables required relatively free productive use of the L2... (p.483)

This possibly fabricated the illusory superiority of prompts in automatizing learners’ retrieval of the form. The use of free production tasks may help orient learners’ focus to meaning and explicitly or implicitly to the form when applying it to the specific meaning. Free production may hence more accurately evaluate learners’ automaticity in retrieving the form, targeted by either explicit or implicit feedback, in their spontaneous production.
Ellis et al. (2006) adopt metalinguistic explanation, classified as one of the prompt types according to Lyster and Ranta’s (1997) model, to give explicit feedback, and recasts to give implicit feedback in order to explore their effectiveness on learners’ explicit and implicit knowledge of the English past tense –ed form. Ellis et al. found that metalinguistic explanation was more effective than recasts in aiding learners to have more control of their already-known –ed past form. Moreover, metalinguistic explanation, though as an explicit feedback, was found facilitative to learners’ implicit knowledge of the form. As argued by Ellis et al. (2006), this was evident in the delayed effect of metalinguistic explanation on learners’ generalization of –ed form to new items in the tests, apart from old items encountered in the treatment.

The above study appears as another support for prompts’ superiority over recasts. However, the same argument of prompts’ frequency variable borne in Ellis et al.’s (2006) operation of metalinguistic explanation may explain the biased result. In Lyster and Ranta’s (1997) categorization of prompts, repetitions are often with intonational emphasis, and metalinguistic explanation points at the nature of errors through metalanguage. Instead of performing metalinguistic explanation exclusively, Ellis et al. (2006) made repetitions precede it. This again created the double prompting effect as in Lyster (2004), Ammar and Spada (2006) and Ammar’s (2008) studies. Hence, it was confusing as to whether the superiority of metalinguistic explanation was due to its inherent effectiveness or the reinforced prompting.

Moreover, the three instrument tests designed by Ellis et al. (2006) seemed to have favoured explicit type of treatment. As critiqued by Norris and Ortega (2000) and discussed by Ellis (2005a), valid measurements of learners’ implicit knowledge need to be meaning-oriented and calling for learners’ feel of the form under online
pressure, rather than memory of the rules under no time pressure. Although Ellis et al. (2006) used an oral imitation test to look into learners’ online production of the form, it carried controlled meaning without eliciting learners’ free production, where meaning focus overrides form. With all the three tests (oral imitation test, grammaticality judgment test and metalinguistic knowledge test) presenting fixed meaning and probably channeling learners’ explicit focus towards form, the explicit form-focused effect of metalinguistic explanation may have been favoured.

Last but not least, the choice of the target form, –ed form past tense, in Ellis et al.’s (2006) study may have caused the failure of recasts. The sociolinguistic investigation of word-final consonant t/d deletion (Bayley, 1996; Smith et al., 2009) suggests that this phonological phenomenon is common across different English dialects. T/d final consonant deletion, or regressive assimilation (Roach, 2009), has been observed most robust when the following phonological environment is a word starting with a consonant, e.g. loved the dog. In light of the t/d deletion feature in English, it is doubtful how the teacher and learners in Ellis et al.’s (2006) study could distinctively detect each other’s utterance of t/d in non-syllabic regular past form, e.g. walked. The implicit feedback given by the teacher’s recasts on learners’ spoken –ed form could have sounded unintelligible, due to the non-syllabic and t/d deletion features of some –ed form. Therefore, explicit metalinguistic explanations may be a more appropriate feedback choice for disambiguating the phonological feature of –ed form. Metalinguistic explanation’s advantage over recasts in targeting the spoken –ed form may have been the reason why it was shown more effective in Ellis et al.’s (2006) study.

Ellis (2007) steps beyond the comparison between the effects of recasts and metalinguistic feedback on the single feature –ed past tense to their comparative
effects on the two features –ed past tense and –er comparative form. He attempts to make a case for the effects of both the feedback type and target form. Recruiting lower intermediate East Asian-dominant adult learners at a private language school of English, Ellis chose to study the –ed and –er forms because they were already present in learners’ production but lacking in their proficient mastery. He therefore views them as amenable to intensive corrective feedback over a short period of time. The –ed form was originally hypothesized to be an easier form to acquire than the –er form because it: 1) only entails the morphological grammar domain but the latter includes both morphological and syntactic domains; 2) is more frequent than the latter as recorded in corpus; 3) does not need to take care of other sentence clauses like the latter, but only attaching the –ed morphology to its own lexical item; 4) is an easy piece of explicit knowledge requiring –ed in base verbs, but the latter is added to adjectives dependent on their syllabic structure; and 5) covers a larger generic scope applying to every regular verb, whereas comparative form includes –er, phrasal (e.g. more) and suppletive forms (e.g. better). Meanwhile, both forms are similarly unreliable, with past tense being formed by both –ed regular and irregular forms, and comparative with the different aforementioned forms. Moreover, both are semantically redundant because past tense can be replaced by temporal adverbials or context, and comparative by the sentence structure.

Regardless of the above comparisons, recasts were found exerting no significantly different effects on both forms from the control group. Although metalinguistic feedback demonstrated significant effects, it facilitated learners’ more significant performance of the –er form in the immediate term and the –ed form in a delayed manner. The results seemed to counter previous findings that recasts fostered L2 learning (e.g. Doughty & Varela, 1998) and hypotheses that the –ed form is
acquired earlier than the -er form. The factors of saliency and frequency may also explain Ellis's (2007) findings. First of all, the recasts in the study were not given as intensively and saliently as Doughty and Varela's. The treatment period lasted only an hour. Recasts were given in a whole meaning-oriented classroom context, and very briefly containing only a single word each time. Second of all, the learners had well developed the -ed form rule more than the -er form at the start. There may have had more room for learners to develop the -er form facilitated by metalinguistic feedback. Moreover, the more frequent feedback input of the -ed form throughout the treatment may have sustained learners’ exposure across the time-taking process of turning their explicit knowledge to implicit knowledge later. Overall, factors of saliency, frequency and learners’ prior knowledge collaboratively contributed to the unexpected results.

Some may argue that comparing two maximally different feedback may not be fair enough to draw any convincing conclusion. For example, recasts and metalinguistic feedback differ on two grounds, explicitness and the elicitation of learner responses. McDonough (2007) eliminated the factor of explicitness to minimize the scope of difference between the two comparative feedback types in her study. She compared the effects of recasts and clarification requests, which are both implicit, on learners’ development of advancing past tense marking to activity verbs. Activity verbs are predicted by Vendler (1967) as the least likely to be past-tense marked compared to achievement and accomplishment verbs, because of their incompatibility with the completeness meaning of simple past. Vendler’s verb classes will be defined in chapter 4. McDonough counted learners’ past tense marking according to the type emergence of activity verbs, the correct grammar form, and appropriate usage of past time. Past tense verbs were chosen because of the ease of eliciting them in communicative tasks. Like Doughty and Varela’s (1998) count of
interlanguage form, McDonough did not consider learners’ accuracy but their emergent formal and functional past marking in activity verbs.

Results showed that recasts and clarification requests did not differ significantly in promoting learners’ emergence of simple past activity verbs. Although clarification requests are prompts and prompts were found better than recasts in aforementioned studies, they functioned similarly to recasts. Their being an implicit prompt may have attenuated their wide difference from recasts. Furthermore, recasts, though not forcing learners’ responses, were shown able to project effectiveness. Where a feedback situates along the explicit-implicit continuum may determine its effectiveness compared to others along the same continuum.

Different quality, timing and characteristic of responses to recasts have been shown affecting learners’ noticing of the corrective purpose and L2 learning involved. Response to recasts is a prominent feature distinguishing recasts from prompts. As classified in Lyster and Ranta’s (1997) classroom study, prompts necessitate learners’ immediate reformulation of their previous errors as hinted by interlocutors; whereas recasts do not require learners’ responses, as reformulations are directly given by interlocutors. The pushed output (Swain, 1985, 2005) element involved in prompts makes them more explicit or salient in drawing learners’ attention to form. This overt advantage of prompts has been shown more effective than recasts in studies discussed above.

On the other hand, there are also findings showing that the discrepancy of prompts and recasts in pushing modified responses does not draw a difference between the two in terms of effectiveness. Lyster and Izquierdo (2009) conducted a controlled study to meticulously investigate the different output effects of prompts and recasts on adult acquisition of L2 French grammatical gender. They predicted that
prompts, which activate learners’ productive rather than receptive process in learning, are more effective than recasts. With learners’ noticing the error, retrieving the target form, and reformulating their non-native utterances, prompts are appreciated as permitting learners’ deeper cognitive processing than the mere imitation verbatim that recasts require from learners. Lyster and Izquierdo further speculated that the deeper processing that prompts invite may assist with learners’ skill transition from possessing declarative to procedural knowledge (DeKeyser, 2007). In contrast to these expectations, the study found that prompts and recasts exhibited similar effects on learners’ use of the form in different oral production tests across time. Moreover, both feedbacks similarly decreased learners’ reaction time in responding to the computerized binary-choice questions on French gender. Learners’ proceduralized use of the form in both the oral production and computerized tests seemed to have been facilitated equally by prompts and recasts.

The reasons discussed by the authors were several: 1) the controlled arrangement of the study made recasts appear intensive, consistent, short, simple with single error change, and as frequent as prompts; 2) the L2 classroom was heavily form-oriented with form-focused instruction held before the feedback treatments, which may have attenuated the respective effects of prompts and recasts; 3) the level of the learners was intermediate and generally assumed form-oriented. The form-oriented characteristic of the classroom and learners and recasts being explicit probably led to the parallel effectiveness: 1) recasts intensively provided positive evidence and therefore guaranteed exposure for learners to infer their negative evidence; and 2) prompts provided both salient negative evidence and opportunities for learners to produce modified output. Overall, Lyster and Izquierdo’s (2009) study enlightens that the effectiveness of recasts may not rely on learners’ following
responses but their explicitness, despite the study’s deliberate effort in controlling learners’ self-repair after prompts and discouraging responses after recasts. The argument that recasts carry different forms and should not be taken absolutely implicit will be discussed in the next section on frequency and saliency.

Apart from the difference in pushing output, prompts and recasts were also found different when targeting different target forms. Yang and Lyster (2010) investigated the different effects that prompts and recasts engender when targeting the regular and irregular past forms in English. Participants in the study were Chinese EFL university students. Although the Chinese learners had been taught past forms, the authors specified that most were only familiar with discrete grammar forms in decontextualized usage. This is mainly due to the form-oriented EFL education in China. In contrast to the traditional form-oriented approach in Chinese EFL classrooms, Yang and Lyster employed an integrative pedagogy by combining form-focused activities and corrective feedback to inspect the different benefits prompts and recasts bring to rule-based and exemplar-based forms. The findings were prompts resulted in the significant increase of learners’ achievement in using both regular and irregular past tense. However, recasts only led to learners’ significant gain in using irregular past tense. These two feedback groups outperformed the control group significantly. An oral narrative production test was used to assess learners’ implicit knowledge through their spontaneous use of past tense (Ellis, 2005a); and a written narrative production test was employed to assess their explicit knowledge, because more thinking time is allowed in the written mode. Word cues were given to assist learners’ production; this may have however defeated the original spontaneity purpose. All in all, prompts appeared most superior across time, mode and linguistic structures among the three groups.
Yang and Lyster attributed the different effects of prompts and recasts caused by the different forms to pushed output, saliency, and Skehan's (1998) model of dual-mode system. Regular past, with low communicative value and voiceless morphology -ed, may have been brought to higher saliency to attract learners' attention by the pushed self-repair in prompts; whereas recasts, with less clear negative evidence, may not have contributed to help learners notice regular past tense. Moreover, Skehan's model explains that rule-based grammar, as in regular past with consistent -ed rule representation, entails learners' computational processing. It will be difficult for learners to process rule-based grammar online, because of the pressing time to generate and assemble the form amid the ongoing task. Prompts, which actively draw the negative evidence of regular past to learners, may accelerate learners' processing of the form online. On the other hand, irregular past, which is exemplar-based grammar with no clear rule, only requires learners' quick retrieval of holistic items. Learners can then learn the form just by hearing it from the positive evidence in recasts. Accordingly, recasts may be more suitable for targeting exemplar-based grammar.

This section notes the difference between prompts and recasts, and discusses the superior effect of prompts over recasts in eliciting learners' self-reformulation. Similar to Lyster (2004), Ammar and Spada (2006), Ellis et al., (2006), Ellis (2007), McDonough (2007), Ammar (2008) and Lyster and Izquierdo (2009), Yang and Lyster’s (2010) results may have been due to the way recasts were being delivered. As discussed previously, the earlier studies may have enhanced the frequency or saliency of the form of feedback, leading to prompts' superiority over recasts. Yang and Lyster also admit that recasts in their study were delivered somehow non-saliently; for example, with multiple error changes and non-corrective repetitions.
occurring in the treatment sessions. The corrective salience of prompts with focused negative evidence and pushed output may have thereby overshadowed that of recasts in the study.

2.7 Frequency and Saliency

The above section shows that frequency of supplying prompts and saliency of turning recasts explicit may have generated their facilitative effects in learners’ L2 learning. The explicitness of prompts is their indicating learners’ noticing of the negative evidence about what is wrong with their non-targetlike utterances, as well as prompting learners’ modified output. These advantages of noticing and pushing output in prompts are correspondingly limitations of recasts (Loewen & Philp, 2006). Loewen and Philp consider that these two limitations of recasts may be resolved by enhancing the degree of contrast recasts make between learners’ non-target utterances and teachers’ target reformulation. The contrast made is important in motivating learners to notice their problematic forms (Saxton, 1997), and map the specific form and meaning for later acquisition (Doughty, 2001). This importance of error-and-target comparability has been argued by Loewen and Philp as a more necessary benefit of recasts than explicitness.

Developed from Loewen’s (2004) study, Loewen and Philp (2006) investigated the different characteristics that would turn recasts more beneficial in driving learners’ successful uptake and subsequent use of the form. Using the same form-oriented ESL context, database and individualized tests, Loewen and colleague observed that recasts were similarly beneficial as other prompt types of feedback. This was likely because recasts in their ESL adult classrooms appeared as a segmented, short, one-change reformulation, declarative statement, stressing the
reformulated form, and being used independently from other corrective feedbacks in the classrooms. These characteristics were found highlighting the corrective contrast and effectiveness of recasts in different ways: 1) learners' successful uptake after recasts was facilitated by characteristics of stress, declarative intonation, one change, and multiple feedback moves; and 2) learners' posttest accuracy was facilitated by characteristics of interrogative intonation, shortened length, and one change.

The results showed that recasts comprising one change promoted both learners' successful uptake and accuracy. Loewen and Philp elaborate that little difference between recasts and learners' erroneous utterances may foster their comparability, and consequently juxtaposing the precise contrast to learners. Other characteristics did not seem to jointly contribute to learners' successful uptake and performance accuracy; they may however unveil their particular functions.

The clearest dissimilarity is declarative intonation versus its interrogative counterpart. The authors reasoned that declarative intonation led to learners' successful uptake, because it conveyed teachers' didactic and form orientation; whereas, interrogative intonation seemed giving learners free choices of either responding to the correction or requesting for meaning confirmation. In view of learners' later performance, the facilitative effect of interrogative intonation was attributed to its choice-giving feature and accordingly learners' deep thinking of their own and teachers' utterances. On the other hand, declarative intonation does not offer choices, and its didactic message may have led to learners' mere repetition entailing little cognitive engagement.

The effects of stress and multiple feedback moves on learners' successful uptake probably came from their help with learners' realising the form in focus; this corroborated earlier classroom studies. Last but not least, the shortened length of
recasts, speculatively when targeting a new form and learners could not confidently produce it immediately but later after further consolidation, only supported learners’ posttest accuracy. Egi (2007a) and Philp (2003) also yielded similar findings concerning the length of recasts. Notwithstanding Egi’s (2007a) insignificant finding of the target form variable, which will be illustrated in section 2.8, the length as well as the number of changes in recasts had been found significantly affecting learners’ interpretation of recasts. Long recasts with more changes to learners’ original utterances were more readily interpreted as targeting content; short recasts with few changes were more likely interpreted as targeting linguistic errors. All in all, the one-change characteristic was found more consistent than other characteristics in sustaining learners’ immediate as well as subsequent production in Loewen and Philp’s study. This may confirm their argument that characteristics shaping the comparability of recasts and learners’ errors are more beneficial than those projecting the explicitness of recasts.

Instead of being categorically placed in either end of an explicit-implicit corrective feedback continuum, Sheen (2006) suggests that recasts run along their own explicit-implicit continuum depending on their different variables. Sheen comprehensively and specifically illustrates variables which contribute to the explicitness and thereupon effectiveness of recasts: 1) phonological focus bearing communicative value in the communicative context; 2) shorter in length than learners’ original utterances; 3) reduced or partial reformulations without repeating learners’ entire utterances; 4) declarative mode presenting reformulation in a statement rather than in any discourse-like ways, for example interrogative confirmation check; 5) interlocutors’ full or partial repetition rather than combined use with other feedback types; 6) single error focus targeting one error at a turn with no other error
reformulations; 7) isolated with no meaning addition; and 8) substitution change replacing one erroneous form with a target form rather than deleting, adding, or reordering original forms. Though Sheen’s (2006) study, like Lyster and Ranta’s (1997) study but more form-oriented or didactic, is a classroom observation relying on learners’ immediate uptake as evidence, she stresses that uptake may only be taken to substantiate learners’ noticing to some extent with no direct link to indicating learning.

The above studies have discussed the different variables conducive to the effectiveness of recasts from the angle of learners’ interpretation and responses. Kim and Han’s (2007) stimulated recall study investigates the teachers’ side in addition, to investigate the extent of learners’ recognition of the gap between their errors and teachers’ recast. Their study found similar results as the above studies from their EFL students: learners mostly recognized the gap 1) when recasts were simple involving no more than one error change, 2) when the linguistic target was not morphological, and 3) when the form of recasts was an isolated declarative statement with no meaning addition.

Kim and Han’s (2007) study possesses another value of probing into two other variables. They are teachers’ intent on giving corrective (form-focused) or communicative (meaning-focused) recasts, to explore if learners’ interpretation overlaps with teachers’ intent; and the directness of recasts given to learners, to see if recasts are better perceived by direct or indirect addressees. The results showed that learners’ interpretation matched the teachers’ corrective and communicative intents for both simple (one change) and complex (more than one change) recasts; and learners perceived recasts equally no matter whether they were directly addressed to learners themselves or their classroom peers.
The above study illuminates that different characteristics of recasts specialize in assisting learners’ successful uptake and subsequent learning in certain enhanced ways. However, to reveal if recasts with these different characteristics are better than those without them in effecting learners’ repair, research is needed to compare the varied forms of the same feedback type. Nassaji’s (2007) non-classroom dyadic task-based interaction study compared the explicit and implicit forms of reformulation and elicitation types of feedback. It explored if explicitness or saliency converts one type of feedback from not generating learners’ output to the opposite. He classifies recasts as the reformulation feedback type, Lyster and Ranta’s (1997) prompts as the elicitation feedback type, and names intonational or verbal signals as prompts.

The results showed that both reformulation and elicitation types of feedback led to higher repair rates when accompanied by salient prompt signals than their less explicit counterparts. Isolated recasts which only reformulate the error form triggered more learners’ successful repair than embedded recasts which reformulate the error form and repeat the rest of the learner’s utterance. Recasts generally motivated more successful repairs than elicitations. This was due to the fact that more recasts with enhancement prompts were delivered than elicitations with enhancement prompts. The study also exhibited more frequent use of recasts and elicitations with prompts than without prompts, and generally more reformulation than elicitation type of feedback. These showed the heavier form orientation of Nassaji’s dyadic task-based context than a naturalistic classroom setting, as well as its preference for the communicative function of recasts to sustain the task flow. Nassaji’s study confirmed the role of explicitness and salience in influencing the repair-generating potential of feedback, even when the feedback does not necessarily push for output.
Nassaji's (2009) later study advanced to investigate the effectiveness of specifically the varied explicit forms of recasts and elicitations versus their implicit forms on learners' subsequent learning. Following Loewen's (2005) practice of measuring the learning effect of unplanned recasts on incidental errors, Nassaji also tailor-made individualized tests asking each learner to identify and correct the same errors they made and being treated by either recasts or elicitations during the interaction task. The difference, also a refinement to Loewen's study, is that Nassaji included a pre-interaction test to closely gauge learners’ acquisition across time. In the pre-interaction stage, learners were first asked to write a description based on a sequence of pictures that they had re-ordered. Then, during the interaction treatment, learners received feedback while orally narrating the picture sequence as similar as possible to what they had written before. Immediately afterwards, they were given back their pre-interaction written account and instructed to identify and correct the errors. The 2-week-later delayed posttest also asked learners to identify and correct the errors based on their same pre-interaction written account. All these steps were designed to trace learners’ learning of their non-slip-of-the-tongue errors in their output through recasts or elicitations over time. Only those pre-interaction errors repeated in the treatment interaction were targeted by feedbacks; this could ensure that the feedbacks were targeting learners’ errors which were no slips of the tongue.

The results of Nassaji's (2009) innovative pretest-posttest design found that recasts facilitated learners’ immediate post-interaction correction more effectively than elicitations; but learners recalled the errors and corrections more from elicitations than recasts in the delayed post-interaction test. Moreover, the explicit forms of both recasts and elicitations led to more post-interaction correction than their implicit forms. However, the explicit form of recasts projected more immediate corrective
effect than the explicit form of elicitations. Nassaji attributed the more pronounced
effectiveness effect of recasts to their targeting only recurrent errors from the pre-
interaction task. Repeated errors as well as errors occurred in both oral and written
modes would least probably be learners’ slips of the tongue on their already-learned
forms. Since recasts entail positive evidence providing reference of new forms
(Leeman, 2003), they would be more effective in targeting learners’ non-performance
mistakes (Corder, 1967). The lesser explicitness effect exhibited by elicitations, which
push learners to self-correct known forms without providing the target form, may
therefore be justified. Moreover, the more sustained effect of elicitations than recasts
may suggest that self-corrected forms could stay in learners’ minds longer than other-
reformulated forms. Overall, the factor of learners’ knowledge of the form determines
the effectiveness of recasts and elicitations when both carry explicitness; and self-repairing helps maintain learners’ learning of the form across a longer time.

The limitation of Nassaji’s (2009) use of individualized tests is also the point
made previously for Loewen’s (2005) study. The point is that gauging learners’
learning through their recall of the exact error and correction may not accurately
probe their extensive control of the form targeted by spontaneous feedback.
Moreover, testing discrete forms may only tap learners’ explicit knowledge; learners’
implicit knowledge, which can be measured through multiple use of the form in
learners’ spontaneous oral output, may also need to be examined (Ellis, 2005a).
However, spontaneous output can contain learners’ performance mistakes (Corder,
1967) instead of or besides their interlanguage errors, which may make feedback
succeed easily and distort results. Loewen (2005), Loewen and Philp (2006) and
Nabei and Swain’s (2002) use of learners’ spontaneous output may have been why
their learners’ correction rates were higher than those of Nassaji’s study.
The use of appropriate tests to gauge learners' learning is important to avoid distorting results. Ellis (2005a) conducted a study to inspect the construct validity of some language tests commonly used to measure L2 learners' proficiency in terms of their explicit and implicit knowledge. He first of all characterizes explicit knowledge as learners' conscious awareness of linguistic features, declarative knowledge of grammar rules, irregular or inconsistent exhibition of knowledge, controlled access to knowledge, accurate access to knowledge, verbalizable knowledge, and knowledge that can be learned at any age via formal instruction. Implicit knowledge is on the other hand the opposite of explicit knowledge in terms of the above characterizations. Though the explicit/implicit and declarative/procedural dichotomies share the feature of rule-based or usage-based, they differ in that declarative/procedural knowledge is about the awareness of the rule of form behind the use of form in communication. One can use a form automatically in communication with or without being aware of its rule behind.

Moreover, Ellis specifies the different indications of learners' explicit and implicit knowledge in different language tests; and exhibits the different degrees of construct validity of these tests in triggering the different indications. The factor analysis illustrated that oral imitation, oral narration, and timed grammaticality judgment tests were valid in measuring learners' implicit knowledge, through eliciting learners' systematic responses according to feel not rules and imposing time pressure to motivate usage automation. For measuring explicit knowledge, untimed grammaticality judgment and metalinguistic knowledge tests were found valid in eliciting learners' variable responses constrained by rules, and allowing time for learners' controlled access to their linguistic knowledge.
Not only is the type of measurement an external factor affecting the effectiveness of recasts on learners’ learning, but also how teachers implement recasts in communicative L2 classrooms. Recasts have been accused of being ambiguous because of their carrying double functions of correcting learners’ errors as well as maintaining the flow of the ongoing conversation (Lyster, 1998a). Long’s (1996) long-standing position of incidental meaning-based recasts and Lyster’s (Lyster & Ranta, 1997; Lyster, 1998a, b; Lyster, 2004) classroom studies have defined and shown the implicitness and ineffectiveness of recasts. Han and Kim (2008) defend the effectiveness of recasts and suggest that their potential success is dependent on teachers’ strategic use in classrooms. They first of all argue why recasts are favoured despite their notorious ambiguity, and then propose five strategies that teachers can employ recasts to enhance learners’ noticing the gap between their errors and the corrected form.

Han and Kim (2008) explain that recasts are favoured and used frequently, though often not systematically, in L2 classrooms because they are congruent with the inherently communicative nature of most L2 classrooms. Moreover, they are contingent upon learners’ meaning, and therefore can be of interest to learners as well as free their attention to the form targeted. The juxtaposition between learners’ errors and teachers’ immediate recasts can also limit learners’ attention to the changes made. However, recasts may not be as effective as expected because of their ambiguity and non-systemic use. Han and Kim recommend the consistent use of recasts in classrooms and teachers to engineer the salience level of recasts: 1) taking advantage of learners’ “natural, perceptual tendency” (p.5) of being more sensitive to simple declarative recasts than complex interrogative recasts and using the former for corrective purpose and the latter for communicative purpose; 2) narrowing the focus
via simple and partial recasts before using simple and full recasts; 3) clarifying learners' meaning first before using meaning-based recasts; 4) motivating learners' repair by first using prompts (Lyster, 2004) to foreground recasts and initiate learners' incorporation of the correction into their output, for example the use of repetitions preceding recasts in Doughty and Varela (1998); and 5) grabbing the golden chance by delivering recasts when learners show signs of needing help with forms, e.g. hesitations or self-repairing attempts.

Not only do feedbacks have their salient and non-salient forms, different English L2 forms can also be distinguished along the saliency scale. However, unlike the former, the latter does not allow turning the same form from its non-salient self to the opposite. In other words, an L2 form can never become salient if it is intrinsically non-salient, for example grammar morphemes. Therefore, to help learners attend to grammar morphemes, feedbacks need to be made salient, focused, and consistent; for example the consistent supply of enhanced recasts targeting predetermined grammar in Doughty and Varela's (1998) study. If not then, recasts, especially unenhanced ones, may be beneficial for salient and meaning-bearing forms rather than non-salient forms (Long, 2007).

This section draws the effect of frequency and saliency in making recasts more effective than they normally are in communicative classrooms. Loewen and Philp (2006), Sheen (2006), and Kim and Han's (2007) studies showed the different ways of making recasts more salient to learners. Nassaji's (2007, 2009) studies particularly showed the effectiveness of explicit recasts over implicit recasts. In practice, Han and Kim (2008) suggest certain ways to deliver recasts explicitly and thereby effectively in classrooms. The external help of raising the salience and consistency of recasts may compensate learners' limited internal capacity
hypothesized by the noticing hypothesis (Schmidt, 2001) and information-processing theory (Skehan, 1998; VanPatten, 1990, 1996, 2004). The salience in recasts may also help counter the cue salience of learners’ L1 in overshadowing their L2 learning (N. Ellis & Sagarra, 2010b).

2.8 Target Form Influence

2.8.1 Interpretation of Form in Recasts

There is some empirical evidence from examining learners’ perception of recasts which supports the view that recasts may be less effective when targeting morphosyntactic errors in general. First of all, Mackey et al. (2000) obtained from their learners’ stimulated recalls that interactional feedbacks to learners’ lexical, semantic and phonological errors were clearly perceived as intended. Feedbacks to learners’ morphosyntactic errors were however perceived as targeting semantic errors for ESL learners and lexical errors for Italian as foreign language learners in the study. The interactional feedbacks naturally arising in individual interviewer-learner pairs’ communicative tasks were recasts, negotiation and combination of both. In addition to learners’ perception of the different feedback targets, Mackey et al. (2000) also found that recasts were mostly used for morphosyntactic errors. They speculate that recasts, which do not push for output, may have been the reason why learners tended not to perceive their morphosyntactic errors. The communicative context may have been another reason why morphosyntactic recasts were not perceived as such. Mackey et al. (2000) suggest that lexical and phonological feedback would contribute more to learners’ understanding of the ongoing communication, but morphosyntactic feedback was comparatively lower in communicative value and hence least probable for learners to notice amid the
meaning-oriented context (VanPatten, 1990, 1996). All in all, Mackey et al.'s (2000) study breaks new ground of revealing the corrective ineffectiveness of morphosyntactic recasts through learners' retrospective perception. Mackey et al. (2000) caution that immediate uptake can only be used to suggest learners' perception of the feedback at that moment.

Carpenter et al.'s (2006) more controlled experimental study corroborates Mackey et al.'s (2000) earlier finding that learners tended not to perceive morphosyntactic recasts as intended. Apart from this shared piece of evidence, Carpenter et al. (2006) ventured an innovative methodology to look into learners' recognition of recasts with or without the juxtaposition of learners' original erroneous utterances. The innovation was using edited video clips of other learners' task output to show to the participants the learners' erased erroneous utterances preceding interlocutors' recasts or repetitions for one group, and non-edited ones with the presence of both the interlocutors and learners' utterances for another group. The results showed that the participants recognised recasts after viewing video with the learners' original errors more than that without learner errors. The percentage of recognising recasts of each group was nevertheless less than 40%. This has been argued as confirming the hypothesis that recasts are better recognized as corrective by learners under the immediate contrast between their errors and the feedback given (Doughty, 2001; Long, 1996, 2007; Saxton, 1997).

However, as Carpenter et al. (2006) admit, their study involves the flaw of detaching participants from the contingent recasting condition by asking them to view other learners' recasting episodes. This may have made them distant from the immediate experience of contingent recasts, and hence interpret most recasts contingent on the immediate speech errors of the videoed learners as non-corrective.
The study controlled participants’ second-hand experience by familiarizing them with the entire process of the instruction and interaction tasks before viewing other learners’ video performance. This was to equalize their content understanding as well as error experience with those learners’ on the tape. Having these controlling measures, Carpenter et al. (2006) therefore attributed the tendency of participants to view recasts as repetitions, to the potential corrective ambiguity of recasts in the content-based discourse context.

The methodological flaw of distancing participants physically, also temporally, from the contingent recast context could still have been the cause of participants’ misinterpreting most recasts as non-corrective. According to Long’s (1996) fundamental definition, L2 recasts are entire or partial reformulations of learners’ incidental errors naturally occurred during ongoing meaning interaction. Lacking the personal participation in exhibiting their natural errors, Carpenter et al.’s (2006) participants may have become less able to empathise the contingent corrective effect of recasts on the videoed learners. Moreover, based on Norris and Ortega (2000) and Ellis’s (2005a) conclusion of the proper measurements of explicit and implicit knowledge, controlled and detached measurements may not favour implicit treatments. The detached use of other learners’ recast episodes to measure participants’ recognition of recasts may not have facilitated the implicit corrective effect of recasts on the participants, though they underwent recasting on the target form in a prior training session. As a result, Carpenter et al.’s (2006) study suggests that methodological and measurement flaws can make recasts opaque.

In contrast to Mackey et al.’s (2000) findings, Egi’s (2007a) study which exclusively probes into learners’ perception of recasts found that learners’ percentages of interpreting the different functions of morphosyntactic recasts were
similar to those of lexical recasts. Recasts have been identified as 1) serving both formal correction and meaning confirmation in interaction (Long, 1996), 2) bearing corrective ambiguity due to the presence of frequent non-corrective repetitions in classroom discourse (Lyster, 1998a), and 3) constituting both positive evidence (Leeman, 2003) and negative evidence (Doughty, 2001; Long, 1996, 2007). Egi accordingly dissects the functions of recasts into 1) responses to content, 2) positive evidence, 3) negative evidence, and 4) both positive and negative evidence. Both recasts targeting morphosyntactic and lexical errors were interpreted by learners as performing the aforementioned functions in similar patterns.

This result seems to depart from Mackey et al.’s (2000) finding that learners had lower tendency to perceive morphosyntactic feedback than lexical ones. However, Egi (2007a) explains that her consistent and frequent use of morphosyntactic recasts targeting two particular forms all along may have raised their saliency level to a level comparable to that of lexical recasts. The variables of frequency and salience, which enhanced learners’ noticing of the form (Schmidt, 1990, 1993, 2001), may have misrepresented the parallel effectiveness of morphosyntactic recasts to lexical recasts. A point often overlooked, as Sheen (2006) concludes from Lyster’s (1998a) finding, is that frequency alone does not ensure saliency. The correctiveness of recasts was obscured by the meaning-oriented classroom discourse despite their most frequent use by teachers. This conclusion may not apply to experimental settings where recasts are often given intensively to the same particular forms consistently (Han, 2002; Mackey & Philp, 1998). From Egi’s explanation, target form may not be a crucial factor affecting the effectiveness of recasts as much as the way they are being implemented.
2.8.2 Form Complexity / Difficulty

Among the different L2 forms that corrective feedback usually targets, grammatical morphemes may be more difficult and acquired later than lexical items because of the bound and less content-bearing features of the former (Goldschneider & DeKeyser, 2001). Even within the same category of grammatical morpheme, some forms are more salient, less difficult, and acquired earlier than others. Goldschneider and DeKeyser identified and tested in their meta-analysis of past grammatical morpheme studies that perceptual salience, semantic complexity, morphophonological regularity, syntactic category, and frequency are the five factors best determining L2 learners' acquisition order of grammatical morphemes. They further argue that these five factors are all about salience. In the authors' analyses of the six grammatical morphemes (progressive -ing, plural -s, possessive -s, articles a, and, the, third-person singular -s, and regular past -ed), the general finding is that the more salient the forms are the easier and earlier learners acquire them.

In the case of perceptual salience, Goldschneider and DeKeyser discuss that the more perceptually distinct a morpheme is from its environment, for example progressive -ing, the easier and earlier learners acquire it. For semantic complexity, the less clear one-to-one form and meaning relationship a morpheme entails, e.g. third person -s entails meanings of “person, number, tense and aspect” (2001, p.34), the more difficult and later learners acquire it. For morphophonological regularity, the less phonologically regular a morpheme is in terms of being affected by its phonological environment, e.g. /t/ and /d/ assimilation in a -ed regular past form to its following consonant (Roach, 2009), the more difficult and later learners acquire it.

For the syntactic category, free morphemes, e.g. articles and lexical morphemes, are
easier and earlier to be acquired by learners. Last but not least, the more frequent a morpheme appears in input, the easier and earlier learners acquire it.

The study of acquisition order helps illustrate which L2 grammar form is more difficult to learn and acquired later than the other. Besides specifying the different factors constructing the acquisition order of different L2 grammatical morphemes, DeKeyser (2005) insightfully defines difficulty of L2 grammar. He argues that the difficulty of L2 grammar concerns the degree of transparency of form-meaning relationships. The importance of certain forms for expressing certain meanings may determine form-meaning transparency. For instance, forms which can easily be replaced by other representations in the discourse or sentence may become redundant and hence draw a less transparent relationship with the meanings they express. This aligns with VanPatten’s (1990, 2004) differentiation of redundant and meaningful forms in determining the ease or difficulty of early L2 development. DeKeyser further dissect L2 difficulty into three components, problems of meaning, form and form-meaning mapping; and explains how these components affect transparency. Novel or abstract meanings, e.g. English articles to article-free languages like Chinese, are problems of meaning; and the complexity of choosing the right morpheme among all possible ones for one meaning relates to problems of form.

Problems of form-meaning mapping involve several different factors: redundancy, optionality and opacity (DeKeyser, 2005). Redundancy concerns whether a form is semantically necessary in a context. According to the Lexical Preference Principle (VanPatten, 1996; Benati, 2005), learners prefer to process lexical items to grammatical morphology sharing the same meaning. Benati (2005) quotes the Chinese learners in his study whose L1 is tenseless and uses temporal adverbs and lexical item le to indicate past time. They were found using temporal adverbs to
unequivocally replace tense and aspect in English for the past time meaning. The past tense form may thereby become redundant, though it carries both meaning and grammatical functions favouring form-meaning mapping (Leow et al, 2008). Similar to problems of form, optionality involves alternations between forms meaning the same. Lastly, opacity is either about different forms used for the same meaning, e.g. both regular and irregular past tense forms used for past time reference, or the same form conveys different meanings, e.g. –s in English for present tense, third-person singular and plural. Same as his earlier claim, DeKeyser considers that salience and frequency in input may assist with increasing form-meaning transparency.

According to N.Ellis and Collins (2009), salience and frequency shape the quality of oral input for assisting L2 acquisition. They classify salience and frequency as aspects of category learning, besides form-meaning mapping significance and reliability. Category learning concerns learners’ acquisition of linguistic constructions from their experience with the perceptual and cognitive characteristics of input. Learners’ exposure facilitated by the different input characteristics may enable them to form categories, generalize, and extend their experience to new forms. N.Ellis and Collins further list factors which determine category learning: input frequency, form salience and perception, prototypicality of meaning, redundancy of form, and reliability of form-meaning mapping.

N.Ellis and Collins split input frequency into token and type frequency. The former is the frequent occurrence of a particular form in input; and the latter is the extensive applicability of a morphological or syntactic form to different words or sentence constructions to form certain meanings. The regular past –ed form would exemplify type frequency because it can join a wide range of lexical items to form the past time meaning. Individual irregular past items would illustrate token frequency,
when their frequent occurrences strengthen their influence in learners' acquisition.

Form salience and perception is about how easily can learners detect the cues of forms for certain meanings. Taking the same example of past tense, the conjugation involved is of lower salience than a lexical item. Both providing cues for temporal meaning, but temporal adverbs may overshadow tense markers and be easily acquired by learners. Redundancy is therefore borne by past tense markers. Concerning form-meaning mapping, the authors posit that the more consistent they are drawn and presented to learners, the easier they acquire the connection between forms and their attached functions. They add that it would be easy for learners to acquire the connection if prototypical forms are displayed. For instance, past tense marking occurs mostly in verbs denoting achievements (Vendler, 1967), because their meanings match—past tense describes completion and achievement verbs convey completion without duration. The frequent occurrence and easy acquisition of prototypical past tense may then help learners form the past tense category and generalize the past tense usage to other non-prototypical forms; for example state verbs denoting states with no perceptual change across time, such as *be*.

The input effect of frequent prototypical past tense marking in achievement verbs is identified as *skewed frequency* (Boyd & Goldberg, 2009). Boyd and Goldberg define skewed frequency as providing a limited number of forms which are typically constrained by certain meaning. In the case of past tense marking, the skewed frequency of marking achievement verbs in input may guide learners to group their meaning similarities and whereby acquire the prototypical use of past tense. Although skewed input is different from type frequency, both contribute to constructional acquisition. Constructions exemplify surface form-meaning mapping and thereby disseminate grammar knowledge. Type frequency fosters category
learning from delivering a wide range of item-specific constructions populated with a particular form. Learners are hypothesized to elaborate on the various input exemplars and form abstract representation to drive novel production. The authors however comment that skewed input may not facilitate productive acquisition because of its prototype restriction, though it allows learners to recognise the meaning constraints of a construction. In the case of -ed regular past exemplifying type frequency, Collins et al. (2009) predict that it will however be less likely acquired by learners than -ing progressive regardless of its frequency. They reason that the -ed morpheme entails less perceptual salience than the -ing morpheme in terms of its obscurity in being heard within a speech stream. The obscurity includes its rarity to be stressed and followed by a pause, and feature as a non-separate syllable except its [-ad] allomorph, which is often minimally articulated. Accordingly, perceptual salience may have to accompany type frequency to exert an input effect in speech.

Back to DeKeyser’s (2005) conception, it may have been narrow in viewing difficulty of L2 grammar as determined by the sole relationship between form and meaning. For example, he may consider that regular past tense –ed is difficult because of its less morphophonological regularity (Goldschneider & DeKeyser, 2001) and opacity (DeKeyser, 2005) in distinguishing itself from its phonological and morphological environment respectively for clear form-meaning mapping. Ellis (2006a) on the other hand takes a relative view in interpreting grammatical difficulty. He views learning as comprising both the implicit and explicit knowledge of the same grammar feature, and labels regular past –ed as easy-to-grasp explicit knowledge because of its clear and available rule of adding the morpheme –ed. He at the same time categorizes regular past –ed as difficult when its accurate use as implicit
knowledge is not guaranteed. In other words, a learner may use past tense wrongly in real time but has sophisticated knowledge of the rule and meaning of past tense.

Rothman (2007) argues that there should be the competence-performance division. He justifies that surface morphological errors of grammar may not reflect learners' underlying syntactic and semantic knowledge of the grammar form. The surface errors may only be learners' performance errors (Corder, 1967, 1981) in real-time use. Rothman quotes the example of Chinese learners of English, who are often competent in L2 syntactic and semantic knowledge regardless of their erroneous performance in surface production. Rothman also suggests that studies should take both meaning interpretive and form productive measures to gauge learners' competence as well as performance to attain a holistic picture of their L2 acquisition.

Not only may different grammar features affect learners' difficulty in L2 learning, but also the effectiveness of different form instructions. Varnosfadrani and Basturkmen's (2009) study suggests the latter; whereas Spada and Tomita's (2010) meta-analysis of relevant past studies illustrates no interaction between the type of instruction and the type of form. Varnosfadrani and Basturkmen employed metalinguistic explanation of learners' erroneous form as the explicit correction and recasts as the implicit correction in their study. It investigated the differential effects between explicit and implicit correction and between their effects on early and later acquired forms. The authors did not use any particular way of defining learners' early and later acquired forms, but synthesized different past studies' findings: 1) definite article (the), irregular past tense and plural 'S' are early acquired forms; and 2) indefinite article (a, an), regular past tense, relative clauses, active and passive voice and third person singular 'S' are later acquired forms. Their control group-free experimental study did not aim at probing learners' acquisition over time. In the
study, each learner was asked to read two passages and then retell the content. Each was corrected during their retelling output according to the correction type assigned. The explicit correction group received immediate explicit correction once errors were produced and delayed explicit correction after the retelling account; whereas the implicit correction group received immediate implicit correction and delayed explicit correction. Learners’ corrected forms were chosen to devise individualized tests (Loewen, 2005) administered about a week later.

The results of the study showed that explicitly corrected learners scored higher than their implicitly corrected counterparts; and explicit correction facilitated early acquired form learning and implicit correction facilitated later acquired form learning. The authors attributed the findings to Schmidt’s (1990, 2001) noticing hypothesis. The explicit correction with metalinguistic feedback and corrected form was shown effective in promoting learners’ awareness, identification of the error-and-target gap, hypothesis testing, and perception of the corrective feedback. The implicit correction of recasts did not benefit learners with these advantages. Moreover, Schmidt’s account of noticing and consciousness also explains why explicit correction benefited early acquired forms for which learners were developmentally ready; the metalinguistic feedback may have been too difficult for forms which learners were less ready for. Though the divided effectiveness of the two types of correction on different forms were clearly shown, cautious conclusion is needed because of the unspecified definition of the complexity of the forms. Furthermore, the double use of explicit correction (immediate and delayed explicit correction) may have maximized the effect of the explicit correction.

Spada and Tomita clearly state the different definitions of form complexity. They first of all classify DeKeyser’s (2005) form-meaning transparency definition of
 complexity or difficulty as a linguistic factor. Psycholinguistic and pedagogical factors are introduced as the other two ways defining complexity. The psycholinguistic perspective considers that difficulty occurs when learners are developmentally unready for the target form; and the pedagogical perspective gives the liberty to teachers to decide which forms are problematic with reference to learners’ production errors. Spada and Tomita’s study, which is about the varying effects of explicit and implicit form instructions depending on the grammar feature, employs Hulstijn and de Graaff’s (1994) linguistic perspective to categorize the different degrees of form complexity. Their linguistic perspective guided the study to which form is more difficult than the other according to how many stages of morphological or syntactic transformation a form has to undergo before attaining the correct form. For example, the WH-question formation involving several stages before obtaining a grammatical WH-question sentence would be more difficult than the regular past form which only entails one stage of adding the –ed morpheme.

The result of Spada and Tomita’s meta-analysis found that explicit form instruction (e.g. explicit rule explanations) projected larger effect sizes than implicit form instruction (e.g. picture description, recasts) when targeting both simple and complex grammar forms over time. Therefore, the type of language feature did not interact with the type of instruction. However, Spada and Tomita’s specific definition of complexity may have fabricated this finding. In other words, their use of a different definition may have yielded a different finding. For example, also using the linguistic perspective, Goldscheider and DeKeyser (2001) may consider regular past –ed as difficult because of its less morphophonological regularity; and DeKeyser (2005) may view regular past –ed as difficult because of its opacity as discussed above.
This section shows that target form may cause intervening effects to the effectiveness of recasts. Mackey et al. (2000) and Carpenter et al.'s (2006) studies illustrated that recasts were less perceived by learners as corrective when targeting morphosyntactic forms. Egi (2007a) however found that learners perceived both morphosyntactic and lexical recasts, due to the frequent delivery of the former. Morphosyntactic forms have been classified as difficult by some morpheme studies. Goldschneider and DeKeyser (2001) and DeKeyser (2005) specify factors determining difficulty of forms. Ellis (2006a) however proposes that certain forms can be easy as explicit knowledge and difficult as implicit knowledge at the same time. Drawing the relationship between difficulty of forms and feedbacks, Varnosfadrani and Basturkmen’s (2009) study found the different effects of feedbacks affected by the difficulty levels of forms to learners; whereas Spada and Tomita (2010) did not find such interaction.

2.9 Conclusion

This chapter has provided an overview and critique of past studies on the theories behind and limitations of recasts in projecting effectiveness to learners’ language learning. Recasts are theoretically motivated by the interaction hypothesis, which suggests that scaffolding learners with corrective feedbacks may help focus their attention to form amid ongoing communication. It compensates the flaws of the input and output hypotheses that learners possess the autonomy to absorb input and focus on form in the face of output production. Several L2 classroom studies exhibiting the effectiveness of the different corrective feedbacks make use of learners’ immediate responses to gauge their uptake of the feedbacks. Learners’ uptake of feedbacks demanding immediate responses can be guaranteed, but that of recasts is
questionable due to the already-given reformulation. Learners’ noticing from taking up recasts is also doubted because of potential mimicking.

With learners’ effortless mimicking and the difficulty to indicate their noticing, learners’ extensive application of the target form in other instances may serve a better signal to their attentive responses to recasts. Prompts prominently excel recasts in generating learners’ immediate attentive responses and hence learning; however studies have shown that factors of frequency and saliency of implementing prompts are the driving forces. How recasts are implemented has also been shown determining whether the form targeted can be perceived by learners. Forms with low communicative value in recasts are less likely perceived by learners; however frequent and salient recasts may facilitate learners’ perception. Different target forms also carry different levels of difficulties determined by their frequency and saliency. Apart from their absolute difficulties, their relative difficulties are governed by learners’ different readiness to the same form. The different formal difficulties may also interact with certain types of feedbacks. This suggests the need to consider more the different external factors guiding the most optimal way of giving and studying corrective feedback in general and recasts specifically.
Chapter 3 Literature Review II - Methodological and EFL Variables of Recasts

3.1 Introduction

In chapter 2, it has been argued that Ammar (2008) and Ellis et al.’s (2006) studies used the wrong type of tests to measure the different effectiveness of prompts and recasts, leading to the biased results. Their use of tests calling for explicit knowledge based on controlled processing may have favoured the relatively explicit prompts over recasts. Moreover, the enhanced prompting effect in Lyster (2004), Ammar and Spada (2006), Ammar (2008), and Ellis et al.’s (2006) studies manufacturing the effectiveness of prompts has brought attention to the optimal use of recasts in reinforcing learning from recasts as input and output practice. Different noticing and perception studies have unveiled the unreliability of indicating learners’ noticing. Matters about using learners’ perception to investigate the different internal components of recasts likely drawing learners’ noticing and the different perception-indicating tools are therefore worth delving into. Second language learners have mostly been discussed, foreign language learners’ readiness to receiving communicative and formal feedbacks may also need to be investigated. Certain ways to turn communicative tasks more form-oriented, speaking tasks more manageable, and learning contexts or settings more favourable for EFL learners to benefit from recasts are also noteworthy. The following sections will address the above queries developed from chapter 2 in depth.

3.2 Test Influence

Following Krashen’s (1985) division of L2 learning into learning and acquisition, Ellis (2006a) adopts the dichotomy of easy or difficult explicit knowledge learning and easy or difficult implicit knowledge acquisition. The dichotomy of
explicit and implicit knowledge has been substantiated by the Principal Components Factor Analysis (Ellis, 2005a, 2006a, 2008). It found that the oral imitation test and timed grammaticality judgment test specifically measured implicit knowledge; and the untimed grammaticality judgment test and metalinguistic knowledge test measured explicit knowledge. Thereby, Ellis (2006a) makes use of the tests to identify the more or less explicit and implicit knowledge of the grammar features in his study; investigates learners’ relative difficulty with the features according to their explicit or implicit knowledge; and probes the extent of implicit and explicit knowledge predicting learners’ L2 proficiency.

Beyond the theoretical determinants of implicit knowledge difficulty (frequency, saliency, functional value, regularity and processability) and those of explicit knowledge difficulty (conceptual clarity and metalanguage), Ellis (2006a) used the four aforementioned tests to tap the relative explicit and implicit knowledge of each grammar feature. He empirically found that regular past -ed for example yielded a high explicit score but relatively low implicit score. However, the scores obtained from the tests may have been derived from learners’ guessing, being swayed, or hinted by the constrained dichotomous and multiple choices the tests provided. As mentioned before, regular past -ed was found easy as explicit knowledge with a transparent rule, and difficult as implicit knowledge in actual use. Its item-based irregular counterpart would accordingly be less easily learnt as explicit knowledge without such a rule. More practically, grammar features in their implicit and explicit knowledge senses were confirmed a central determinant of L2 proficiency. Under the use of IELTS measuring learners’ proficiency, the oral part of IELTS was found obligating implicit knowledge, and written IELTS calling explicit knowledge. This determining factor of modality is built on the correspondence between the automatic
processing in oral output and implicit knowledge, and the controlled processing in written output and explicit knowledge.

To conclusively state what learners have learned, valid tests for collecting relevant L2 data are desirable. Ellis (2005a, 2006a, 2008) proposes that what learners have learned concerns the dichotomy of implicit and explicit knowledge; and he (2005a) has illustrated the construct validity of the different tests in measuring learners’ implicit and explicit knowledge. Ellis (2008) later contends that tests measuring implicit knowledge or learners’ interlanguage development, which is considered primary to L2 communicative usage, should tap what learners know as well because it is a building block of what learners can do with the L2. He therefore suggests that a specially designed experimental test which elicits learners’ spontaneous data of using particular forms may insightfully measure what they know. Learners’ production will thereby not be constrained in the same way as an oral imitation test, grammar judgment test or metalinguistic knowledge test.

*Experimental data elicitation* may motivate learners’ incorporation of certain forms demanded into their own production. It may excel *natural production data* in guaranteeing learners’ sufficiently accurate production of the form. Naturally occurring data may however be more convincing because of tapping learners’ free constructed response (Norris and Ortega, 2000). Ellis (2008) thus encourages the use of *oral imitation test*, which does not limit learners’ responses to preset choices but sufficiently taps learners’ self-initiated correction of the form. Nevertheless, learners’ responses were limited to the exemplars provided in Ellis’s oral imitation test. Balancing the weight of measuring what learners know and what learners can do with the L2 may be a valid way to collect comprehensive data about learners’ L2 learning. Ellis’s (2005a) another test of implicit knowledge, *oral narration test*, may fulfil such
a balance. It may elicit learners’ sufficient and obligatory use of certain forms to construct a story meaning, for example using past tense to narrate past events completed one after another.

Apart from the relation between learners’ L2 production and tests, the relation between the effects of different oral corrective feedback and tests is also worth examining. Corrective feedback is a usual facilitator of learners’ L2 production displayed in tests. Loewen and Nabei (2007) inspected if corrective feedback led to learners’ increased performance of English question formation on an oral production task, a timed and untimed grammaticality judgment tests. They further compared the effectiveness of recasts, metalinguistic feedback, and clarification request in advancing learners’ question formation on the three tests. The Japanese EFL students, accustomed to institutional form-oriented instruction of English, participated in communicative tasks with their native-speaking interlocutors. Question forms were chosen because they were the students’ problem likely to trigger corrective feedback; and it is easy to elicit questions in communicative tasks. Under the pretest-treatment-posttest quasi-experimental design, the students received feedback and accomplished the tasks in groups.

The results found were threefold. First, no significant improvement across time was shown on the untimed grammaticality judgment and oral production tests; but significant improvement across time was shown for all groups on the timed grammaticality judgment test, with the feedback groups progressed more than the no feedback group. Second, no significant difference among the three different feedback groups and no interaction between feedback and item grammaticality were shown in the tests. Third, students gained significantly on ungrammatical items of the untimed
grammaticality judgment test and on grammatical items of the timed grammaticality judgment test.

Justifications for the findings can be summarized into the particular operation of the feedback options, more individualized attention on learners, treatment length, contingent feedback moves, construct validity of the tests, and the different measures of learners' production. The insignificant learners' progress on the untimed grammaticality judgment test under all feedback conditions could have been due to the deficiency of feedbacks in influencing learners' explicit knowledge. Instead of discerning what errors learners committed in particular instances, the metalinguistic feedback, though it was the most explicit, only generally informed learners of question formation problems. With metalinguistic feedback being less explicit and recasts being more explicit, the differences among the three feedbacks could have therefore been attenuated. Recasts turned more explicit possibly because working in small groups instead of a large class may have somewhat individualized the attention on learners. Running contrast to their form-oriented large-class EFL practice, small-group meaning-based practice may have appeared exceptional and made recasts more obvious (counter-balance hypothesis, Lyster & Mori, 2006). Moreover, the short treatment length of the study may not have allowed the different feedbacks to significantly exert their different potentials, especially when the oral production test used time-consuming developmental stages as the measure of learners' progress. Since feedbacks were contingent on learners' errors in meaning-based communicative tasks, the feedback moves may not have been sufficient and consistent to project any significant effects.

Regarding grammaticality, the higher score of ungrammatical items on the untimed grammaticality judgment test and of grammatical items on the timed
grammaticality judgment test exactly supported the construct validity of each test, in providing optimal opportunities for learners’ controlled and automatic access to their explicit and implicit knowledge respectively. As Ellis (2005a) found in his analyses of the different L2 grammaticality tests, the untimed one associated more with explicit knowledge in needing time to judge ungrammatical forms and the timed one with implicit knowledge in feeling grammatical forms. However, the oral production test, same as the timed grammaticality judgment test measuring learners’ implicit knowledge, did not show learners’ significant progress. This could have been due to the different measures of L2 production. The use of developmental stages in the oral test may have been less meticulous than the use of discrete accuracy in the judgment test to show learners’ significant changes. All in all, feedback implementation, context, test and L2 measure are possible blocks to a fair view of results. Notwithstanding these blocks, corrective feedback was shown facilitative to learners’ implicit knowledge on the timed grammaticality judgment test in the study.

This section has expounded the different kinds of knowledge that different tests measure. Ellis (2005a, 2006a, 2008) suggests that certain tests are suitable for measuring learners’ explicit knowledge and others for measuring learners’ implicit knowledge. Loewen and Nabei’s (2007) study specifically draws a relationship between corrective feedbacks and tests. They found the different effectiveness of corrective feedbacks in different tests. The different ways of delivering corrective feedbacks were concluded as the main reason behind.

3.3 Role and Optimal Use of Feedback

Although both Krashen (1985) and Ellis (2005a, 2006a, 2008) uphold the dichotomy of explicit and implicit knowledge to explain L2 learning, Krashen
supports a definite non-interface position between explicit and implicit knowledge and Ellis accommodates both the interface and weak interface positions. Krashen radically proposes that L2 is learnt implicitly through exposure to meaning without conscious focus on form. Even when conscious focus on form is in place, it is limited to monitoring learners’ output. Different from Krashen’s view that explicit knowledge will never be transferred to implicit knowledge, Ellis (2006b) recognises explicit knowledge as a means to facilitate subsequent implicit knowledge acquisition.

Ellis (2006b) delineates what, when and how to teach L2 grammar with reference to second language acquisition perspectives. He considers that learners may learn their L2 either along the deductive sequence of rule presentations and then practise in meaningful contexts, or the inductive sequence of deriving the rule on their own after exposing to grammatical exemplars. Any one of these sequences is hypothesized to turn explicit knowledge to implicit knowledge. Meanwhile, Ellis holds a weak interface position that learners may be able to turn their explicit knowledge to implicit knowledge when they are developmentally ready and consciously drawn to attend to the target form. Intensive corrective feedback on learners’ own problematic forms and concentrated focus on single grammar feature may help address learners’ problems directly and stimulate their noticing.

Like Ellis, Dabaghi (2006) sees the explicit manner of error correction as more effective than its implicit counterpart. Although Dabaghi’s study is a short report, it offers rather precise findings regarding the timing, type of features and learners’ developmental readiness factors, which may affect the effectiveness of error correction. Immediate and delayed feedbacks were insignificantly different from each other. This was probably due to an imbalance of feedback types, with explicit feedback (metalinguistic information) given both immediately following errors and in
a delay, but implicit feedback (recasts) was only given immediately. The explicitness factor may have also been a mediating factor. Correction of morphological features was found more effective than that of syntactic features. This was explained as the former is item-based or exemplar-based, and the latter is system-based or rule-based in Skehan’s (1998) dual-mode model. Lastly, explicit feedback was more effective in targeting developmentally early features and implicit feedback more effective for late features. This was because metalinguistic information on late features may have been relatively difficult for learners to understand.

At another end of the interface continuum, N.Ellis (2005) contends that there is a strong interaction between explicit and implicit learning; and learners’ initial explicit knowledge may facilitate their subsequent implicit knowledge after frequent and salient exposure and practice. N.Ellis views learners’ L2 learning process from two cognitive perspectives: 1) from learners’ initial exposure to declarative statement or metalinguistic information of grammar to their creative constructions of utterances, then from their construction exemplars to formula formation, next from formulas to regularity abstraction, and then from abstraction to learners’ generalized usage in other instances, exhibiting implicit learning; 2) from learner’ exposure to exemplars to their later dissection or analysis of the phonological whole to its structural components, next from learners’ conscious rehearsal of the components in the phonological loop of their working memory (Gathercole & Baddeley, 1993) to the association of the discontinuous components, and then from the association between components to regularity abstraction, leading to automatic usage in new instances and thereby illustrating implicit learning.

Any flawed output of learners during the above two processes may be remedied by feedback such as recasts, to bring “a new wave of explicit analysis”
(N.Ellis, 2005, p.28) to learners with target exemplars provided. Immediate feedback is recommended, to avoid interrupting distraction in weakening the strength of learners’ focused attention. Not only can the target exemplar input of recasts provide learners with explicit analysis, but also can learners’ immediate uptake of recasts. Learners’ rehearsal of the target exemplars in recasts in their phonological loop, when receiving the recast or taking up the recast, may offer them conscious analysis of the components of the exemplar, as in process 2) above. In either cognitive view of learners’ L2 learning above, regularity abstraction seems to be a prerequisite for subsequent automatic usage of the form in extensive instances. The extensive use of regular grammar forms or constructions, e.g. regular past -ed, may need the regularity abstraction process. The extensive use of quasi-regular grammar, e.g. irregular past, may need regularity abstraction of its highly frequent items.

Factors of frequency and salience have been repeatedly stated by N.Ellis (2005, 2006) as crucial in enhancing the interface between learners’ explicit learning and implicit learning. They are especially key when there are concurrent multiple cues overshadowing less salient forms in terms of communicative value, phonological environment, and L1 transfer (N.Ellis, 2006; N.Ellis & Sagarra, 2010a, b). N.Ellis quotes the example of regular past tense, which involves non-salient morphological marking. The /t, d/ phonological deletion (Wolfram, 1985; Bayley, 1996), lenition (N.Ellis, 2006), or assimilation (Roach, 2009) in -ed regular past preceding a consonant may be overshadowed by lexically-based temporal adverbials within the same speech stream. He also warns that this overshadowing effect would be worsened, if learners have developed entrenched L1 regularity abstraction of using temporal adverbials instead of past tense marking.
N. Ellis (2005) suggests that not only can frequent and salient exposure to grammar input offer large enough sample to facilitate or speed up the interface between explicit and implicit learning, but also can frequent output practice. The *skill acquisition approach* posits that learning a language resembles acquiring other kinds of skills (Johnson & Jackson, 2006; Scheffler, 2008). The resemblance lies in their learning process: 1) receiving explicit declarative representations in input; 2) retrieving the rules in application, and accumulating enough capacities via rehearsal or practice; and 3) making autonomous real-time usage, with free cognitive effort to engage in other simultaneous domains (Scheffler, 2008). Speaking is an example of skilled behaviour (Skehan, 1998, p.18), where deliberate learning of form may progress to automatic usage of form to convey meaning through form-meaning mapping practices. Form-meaning mapping practices can be realised by tasks making meaning clear enough for learners to channel their focus to the specific form and connect it with its corresponding meaning (Leeman, 2007). Moreover, Johnson and Jackson (2006) contend that skills often require highly proceduralized forms of knowledge, so learner performance-based training or feedback is a recommended way to teach and learn skills.

Pedagogically, language skill can be developed through applying the form in *difficult real operating conditions* (Johnson, 1996), which call for the target form usage in real situations. Through repeated practices (DeKeyser, 2003), learners’ new encounter of declarative knowledge can be automaticized and their limited capacity may thereby be freed (Skehan, 1998). Learners may then be able to realise Anderson’s (1983) early hypothesis that their declarative knowledge (*knowledge about* in Johnson’s term) can be transformed into procedural knowledge (*knowledge how to* in Johnson’s term) assumed in any real-time skill operation. Johnson and
Jackson (2006) add that skill trainers need to address learners’ *performance* or skill-based mistakes, not *competence* or knowledge-based errors (Corder, 1967), to intervene and help learners develop the actual skill beyond their underlying knowledge. Trainers may let learners take some risk in making initial free production, and then provide real-time or follow-up feedback on their immediate performance. This may thereby confront learners with the mismatch between their mistakes and the models. The risk-taking element corresponds to Swain’s (1985, 2005) pushed output and hypothesis-testing theories. The skill approach may be useful to learners who have problematic language use despite their sophisticated knowledge foundation (Rothman, 2007).

Explicit consolidation or reminder of declarative representations via feedback may also be useful to foreign language learners. Their knowledge base may not be sophisticated enough to be self-responsible for analyzing target language codes to build their underlying competence in tasks (Swan, 2005; Scheffler, 2008). Scheffler (2008) argues that language teaching approaches, such as traditional instruction, fixed developmental order, and task-based instruction, used to assemble learners’ L2 knowledge base over-estimate learners’ readiness. Traditional instruction predetermines a structural syllabus which may not align with learners’ internal syllabus. Learners are assumed to follow a fixed developmental order as their internal syllabus; but there may be other factors rather than the developmental order itself that determines which form is acquired first, for example formal salience (Goldschneider & DeKeyser, 2001). Task-based instruction leaves learners alone to experience or experiment the use of the form in the assigned communicative task, and gradually develop the form usage for the particular task demand. However, learners may not be ready for the pre-determined traditional instruction for certain forms because they...
may not be salient to them in certain contexts, nor for analysing the form on their own while doing the task. This is especially the case for foreign language learners who have limited exposure to the target language inside and outside classrooms (Swan, 2005).

Therefore, Scheffler (2008, 2011) recommends that there should be different approaches for first/second language learners and foreign language learners. He suggests that foreign language learners receive explicit and repeated form exposure to facilitate their efficiency of form analysis. This can be done by providing consistent explicit feedback to stimulate learners’ problem-solving cognitive mechanism. Directly and consistently scaffolding learners with using the target form required by a particular task through teacher-led feedback was found by Toth (2011) to have excelled learner-led interaction in providing procedural assistance to learners’ focused use of form.

This reliance on the linguistic environment has been established by the connectionist model. Leeman (2007) recalls that the connectionist model sees linguistic knowledge as developed by drawing associations among forms; and repeated exposure would strengthen the associations. Moreover, when the associations do not yield correct results, their strength may be adjusted by feedback. MacWhinney’s (1987) competition model, a connectionist framework, hypothesizes that forms enter competition with other forms in associating with meanings. Feedback providing positive evidence for one form-meaning association is in turn a negative evidence for other forms which compete for the meaning. In other words, the stronger input cues by feedback override other weaker cues. Stronger cues may also come from learners’ earlier learned forms blocking their later learned forms; and learners’
entrenched L1 overshadowing their later acquired L2 resulting in L1 transfer may exemplify that (N. Ellis & Sagarra, 2010a, b).

Apart from being useful in the connectionist model, feedback can also be important in skill acquisition (Leeman, 2007). Leeman explains that feedback can initially encourage learners' acquisition of declarative knowledge, then remind them to attend to declarative knowledge during the automatizing or fine-tuning process in proceduralization, as well as avoid their automatized use of non-target form. With continuous feedback in output practices, learners' errors may be diminished and less feedback is required. As Scheffler (2008) cautions, learners may lose their declarative knowledge after they have automatized their language skill and are no longer aware of the underlying rules for other-context generalization. This may lead to Selinker's (1972) early conception of fossilization, where learners reach a plateau of form usage and errors cannot be easily rectified. Scheffler thus conceives that explicit instruction or repeated formal feedback is indispensable during learners' production.

The information-processing theory emphasized by Skehan (1998) exactly views feedback as necessary to free learners' cognitive burden in linguistic production with multiple demands—form-meaning relationship, L2 performance accuracy, complexity, and fluency. Feedback to form may help release learners' cognitive load from attention to form to attention to higher-order conceptual processing, or vice versa, to accomplish the production task. Continuous feedback may also facilitate learners' greater attention to subsequent L2 input entailing formal aspects, be it subsequent formal feedback, interaction, or instruction.

This section has viewed the role and use of feedback in effecting learners' learning from the domains of input exposure and output performance. Both Ellis (2006b) and N. Ellis (2005) consider that explicit knowledge facilitates learners'
subsequent development of implicit knowledge. N. Ellis especially suggests that frequent exposure to salient input may facilitate the development of implicit knowledge; and feedbacks may help correct learners’ wrong use of implicit knowledge. Scheffler (2008) on the other hand sees that giving explicit feedback to EFL learners’ performance errors may help remind them of their declarative knowledge, to destabilize their automatized L1.

3.4 Learners’ Perception

Leeman (2007) concurs with previous L2 classroom researchers (e.g. Lyster & Ranta, 1997; Loewen, 2005) that immediate response to oral feedback may not equate learners’ noticing and learning of the form. She notes that learners’ noticing may be more insightfully unveiled by exploring learners’ perception, and their L2 learning may be accurately measured through developmental methods by pre and post-tests across time. Egi (2007b) identifies probing into learners’ interpretation of recasts as crucial in informing which component of recasts contributes to L2 development. She appreciates learners’ perception of recasts as revealing the effect of the different internal multi-dimensions of recasts, adding value over other studies examining the effect of externally manipulating recasts.

As shown in the section about contriving certain recast conditions, recasts delivered with stress, intonation, reduction, and repetition for instance facilitated learners’ noticing of the correction and subsequent learning. However, the different effects of the internal components of recasts have not been investigated widely, especially through learners’ inner thoughts. Recasts are multi-dimensional in terms of providing negative evidence (Doughty, 2001; Long, 2007), positive evidence (Leeman, 2003), as well as content confirmation in conversational interaction (Lyster,
Leeman (2003) is one significant study which dissects the effectiveness of recasts according to their negative and positive evidence functions. Among the groups of implicit negative evidence, enhanced positive evidence, recasts and control, only recasts and phonologically stressed positive evidence learners outperformed control learners receiving models. She therefore attributed the effectiveness of recasts more to their positive than negative evidence. However, the results could have been caused by the additional stress that the positive evidence entailed, and confounded by how the enhanced positive evidence represented the not necessarily enhanced positive evidence element in recasts.

Learners in Egi’s (2007b) study participated in communicative tasks and received recasts by interlocutors. During the treatments, learners expressed their thoughts immediately after each recast instance via immediate cued reports. After the post-treatment immediate post-test, learners were prompted to tell what they were thinking at the previous feedback moment via stimulated recalls. A delayed post-test was held two weeks afterwards. Analyses divided learners’ thoughts into interpreting recasts as responses to content, positive evidence, negative evidence, and negative plus positive evidence. Results first of all illustrated that learners displayed significantly better short-term learning when interpreting recasts as positive evidence alone and a combination of positive and negative evidence than interpreting recasts as responses to content. However, the negative evidence interpretation did not significantly outperform the content interpretation. Moreover, the original hypothesis that the positive evidence interpretation of recasts would significantly lead to learners’ better performance than the negative evidence interpretation was refuted, although this happened with lexical targets.
Evidently, Egi’s findings showed learners’ greatest improvement when interpreting recasts as both positive and negative evidence in both the immediate and delayed post-tests. This was different from Leeman’s (2003) exclusive finding of positive evidence benefit. Apart from the effect of learners’ interpretation, Egi also correlated learners’ learning with the effect of linguistic target. Learners’ interpretation of recasts as positive evidence seemed to facilitate their learning of lexical items; and their interpretation of recasts as both positive and negative evidence led to morphosyntactic learning. She therefore concluded that positive evidence in recasts favours lexical item-based learning from exemplar memorization; and the positive and negative evidence combination favours morphosyntactic rule-based learning from hypothesis testing and rule generalization. Nevertheless, the results need to be taken cautiously because of the use of the problematic tailor-made tests based on learners’ same error items, and the occasions when learners not necessarily verbalize what they notice.

Together with earlier studies (Mackey et al., 2000; Carpenter et al., 2006; Egi, 2007a) introduced in section 2.8.1 of chapter 2, Egi (2007b) probed into learners’ perception to view their inner thoughts about L2 learning. Kim and Han’s (2007) study, which has been discussed in section 2.7 of chapter 2, takes a broader approach than these studies by using stimulated recalls to compare learners’ interpretation with teachers’ intent. Nevertheless, as Kim and Han admit, their implementation of stimulated recall in different time with different learners may have corrupted some learners’ memory of their previous recast episodes. Moreover, the approach of investigating teachers’ intent via stimulated recall, though more comprehensive, may easily incur validity problems. This is based on the conjecture that native-speaking teachers are more cognitively capable than learners; and may therefore have a greater
propensity to learn from the recall process, give expected answers, or evaluate their feedback episodes at the time of recall.

Speaking of stimulated recall, Gass and Mackey (2000) recommend it as one of the methods tapping learners’ noticing of form or their cognitive processes during oral interaction or feedback treatment. It is of a retrospective kind asking learners to verbalize what they were thinking at the time of the activity after its completion. Apart from retrospective types, there are some online types serving the same purpose of tapping learners’ form noticing but doing so during the activity. Learners’ uptake is a spontaneous way to suggest learners’ noticing of form during interaction. However, as noted before, uptake or no uptake may not signal noticing or not due to possible mimicking or conversational redundancy. Bao et al. (2011) even empirically found the superior effectiveness of stimulated recalls over uptake in capturing learners’ noticing, attributing it to the ambiguity and unreliability of uptake. Another online noticing measure is immediate recall. Philp (2003) employed it to primarily measure learners’ noticing the gap between their errors and interlocutors’ reformulated forms immediately after interlocutors’ feedback.

Philp (2003) adopts Robinson’s (1995, 2001a, 2003) definition of noticing, learners’ detection of information in an auditory activated state during short-term memory rehearsal, as the rationale behind her use of immediate recall to measure learners’ noticing of the gap. The proximity of learners’ immediate recall to their errors and interlocutors’ feedback may allow the use of immediate recall to instantaneously capture learners’ activated state of processing the reformulated form just received. The activated state of processing is hypothesized to have followed some allocation of attentional resources within learners’ short-term memory, before permitting information to proceed to learners’ long-term memory. In Schmidt’s
(2001) words, the use of immediate recall is appropriate enough to tap learners’ noticing because “[t]he clearest evidence that something has... been consciously perceived or noticed is a concurrent verbal report, since nothing can be verbally reported other than the current contents of awareness” (p. 20). The use of verbal recall as evidence of noticing can also be justified by Baddeley and Logie’s (1999) multiple-component model of working memory. Working memory is proposed to consist of a system called the central executive; it manages its two subordinate systems called the phonological loop and visuospatial sketchpad. Any audio input (interlocutors’ feedback) is hypothesized as temporarily stored and remained activated in the phonological loop, ready for verbal output (immediate recall).

Though verbal recall may be a valid measurement tool of learners’ noticing of input and particularly the mismatch between learners’ errors and interlocutors’ reformulation in the case of recasts, the measurement of noticing via learners’ verbal recall may not secure internal validity. Schmidt (1990, 2001) and Robinson (2003) recognise the use of verbal recall as one way to measure learners’ noticing; meanwhile, they note that learners’ immediate verbal recall of input may only show their momentary noticing instead of deep enough processing to guarantee retention and later retrieval from memory. The indication of learners’ momentary noticing instead of deep processing may therefore impair the validity of using immediate recall as the measurement tool of noticing.

The way immediate verbal recall is being prompted may also make learners’ evidence of noticing unclear. Take Philp’s (2003) study as an example, she implemented immediate recall through learners’ immediate repetition of the preceding native-speaking interlocutor’s recast, after the interlocutor’s gesture of two knocks on the table to signal learners to repeat. The pre-task training session which modeled the
sequence of the native-speaker’s utterance, then the two knocks on the table, and then learners’ required repetition of the native-speaker’s utterance may have heightened learners’ noticing and retention of the native-speaker’s utterance. Thereby, learners’ successful immediate recall may only be an artifact deliberately manufactured by its implementation as well as the priming effect of the two knocks, rather than a corollary of the effectiveness of recasts in drawing learners’ attention.

Moreover, as Gass (2003) equates learners’ immediate repetition and pure mimicking, Philp’s (2003) use of immediate recall which asked learners to immediately repeat interlocutors’ previous reformulation may lend little support to revealing learners’ deep noticing of the mismatch between their errors and the interlocutor’s reformulation. Noticing the formal mismatch provided by recasts has been hypothesized by Long (1996, 2007) as evidence of learning. Stimulated recall may be a more appropriate indicator of learners’ deep noticing than immediate recall, based on its nature of asking learners to retrospect instead of momentarily mimicking. Nevertheless, stimulated recall possesses drawbacks of memory decay and response fabrication due to its retrospective nature.

This section has discussed the methodology of probing learners’ perception of recasts. Leeman (2007) and Kim and Han (2007) favour the use of perception to investigate learners’ noticing; and Egi (2007b) adds that learners’ perception can be used to unveil the effectiveness of the different internal components of recasts. Bao et al. (2011) even found the superior effect of stimulated recalls over uptakes in probing learners’ noticing. Among all the methods of investigating learners’ perception, Gass and Mackey’s (2000) use of stimulated recall and Philp’s (2003) use of immediate recall have been evaluated.
3.5 Learners’ Level / Developmental Readiness

Notwithstanding the doubtful validity of Philp’s (2003) measurement of noticing, Philp’s findings provide some empirical evidence for the confounding variables of recasts. Learners’ level in terms of their developmental readiness for the target form, length of recasts based on the absolute length of more or less than five morphemes, and number of changes made by recasts were identified as variables possibly constraining learners’ noticing of the mismatch between their errors and the target reformulation introduced by recasts. Mackey and Philp (1998), Philp (2003), Ammar and Spada (2006) and Ammar’s (2008) empirical evidence illustrated in the previous chapter and section comparably suggests that learners’ lower proficiency level or deficient developmental readiness is a candidate influencing learners’ noticing recasts and benefiting from recasts in their learning.

In her empirical study of the effect of recasts on the use of past tense, Han (2002) identifies that learners’ developmental readiness is one of the determining factors. She explains that learners who are developmentally ready for the target form would be more able to notice the implicit feedback in recasts. In other words, recasts are predicted to work better when used to target learners’ already-learned forms. On the other hand, Lyster and Mori (2006) distinguish recasts from prompts by stating that the former is effective in introducing or targeting new forms due to their modeling element; whereas the latter is effective in targeting learners’ already-learned forms due to their quality of eliciting or demanding learners’ self-reformulation. Yet, these researchers interpret whether a form is learned or not based on absolute readiness.

Anderson’s (1983) classification of declarative knowledge and procedural knowledge in his *adaptive control of thought model (ACT)* may introduce relative...
readiness. According to the model, learners undergo three stages before realising skill acquisition: 1) the declarative stage, where rules are learned and can be verbalised; 2) the associative stage, where declarative knowledge are transferred to automatic processing through practice; and 3) the procedural stage, where declarative knowledge is proceduralised and performed without necessarily being able to verbalise the rules behind. Based on these three stages, L2 learners who possess the declarative knowledge of a linguistic form (stage 1), but do not yet possess the procedural knowledge of the form (stage 3), may be ready for the form when stage 1 is considered. They may however be unready when stage 3 is concerned.

EFL learners often have the above imbalance of knowledge acquisition, mainly because their EFL instruction does not incorporate learners’ non-linear language development into syllabus designs. Robinson (2001a) delineates different syllabuses assuming learners’ linear and non-linear language development. According to him, synthetic syllabuses focus on teaching learners language elements in a predetermined sequence, separating learners’ learning and use of language elements by teaching them the easiest or most learnable forms before the harder ones. This kind of syllabus asserts that learners will eventually be able to integrate the use of the separate elements into their real world language performance. Instead of presuming learners’ straightforward language path and disregarding their different formal restructuring processes caused by developmental constraints, analytic syllabuses take a less restrictive approach by providing opportunities for learners to attend to different forms and use them jointly as the communicative activities call for. This proposition attempts to align the syllabus with learners’ internal syllabus or non-linear interlanguage development.
Acquiring communicative skills in L2 may be learners’ goal of meeting real world demands to enjoy what adult native speakers possess (Skehan, 1998; Skehan & Foster, 2001). However, as Skehan and Foster note, native speakers and learners often resort to communicative strategies and prioritize meaning, leading to elliptical or lexical-based communication. The reason for native speakers is claimed to be brevity, and that for learners is their limited attentional capacity in simultaneously processing and producing meaning and form. Hence, Skehan and Foster suggest that continuously engaging in actual communication may contribute short term but not long term benefits to learners. If learners consistently de-emphasize form, their interlanguage change may be inactivated and result in development stabilization (Selinker, 1972; Ellis, 2009a).

Apart from their cognitive constraint causing their form de-emphasis in communication, learners’ insensitiveness in manoeuvering their attention sufficiently to formal input, as neglected by Krashen (1985), and initiating syntactic processing to fine-tune output before or after pertinent feedback, as overlooked by Swain’s (1985) output hypothesis, may also induce learners’ form-focus avoidance. Swan (2005) pinpoints that these cognitive sensitiveness problems are derived from learners’ deficient level of deploying well-equipped existing knowledge in communicative tasks. EFL learners’ constraining learning environment further impairs their approaching advanced or native-like proficiency. Swan (2005) therefore boldly condemns the relevance of the task-based communicative approach to EFL learners, who receive time-limiting classroom instruction and exposure-poor opportunities in and outside class.

Pedagogical irrelevance aside, Swan (2005) presents a theoretical problem behind task-based instruction (TBI), to highlight the importance of proactive formal
intervention to establish EFL learners’ adequate foundation for communicative task performance. He borrows Doughty’s (2001) claim to argue for the flawed belief of TBI in online communication-based acquisition. Doughty contends that occasional shift to focus on form (Long & Robinson, 1998) during online communication may not be adequate, because of the brief availability of learners’ cognitive window to attend to focus on form. Doughty (2001) and colleague (Doughty & Varela, 1998) alternatively propose the use of planned focus on form, which predetermines the target form and targets it consistently, to defeat the ephemeral nature of online focus on form in TBI. The two advocates of learners’ limited capacity, Skehan (1998) and VanPatten (1996, 2004), also advise that learners’ attention to form needs to be driven in some way in order to enhance their focus on the required forms used for certain meanings.

Swan’s (2005) position on the use of form-oriented instruction to strengthen foreign language learners’ grammar foundation is undeniable, because of their lack of target language exposure inside and outside the classroom. However, his radical refutation of task-based learning has been described as unfounded by Ellis (2009a). Swan contends that task-based learning discourages grammar learning and teaching, because of learners’ limited linguistic resources to communicate. However, Ellis defends that this constitutes a misunderstanding of task-based learning and teaching.

Firstly, Ellis (2009a) accommodates that tasks can be both meaning and form-oriented, through contriving the task-meaning context for learners’ use of specific form. This kind of focused task can not only direct learners’ attention to the form needed to fill the meaning gap, but also keep the ultimate focus on conveying meaning by not informing learners of the form targeted. Skehan’s (1998) understanding of task-based learning is however an unfocused one which does not
elicit any particular form. Secondly, Ellis holistically considers that teachers’ explicit pre-task or post-task form instruction and preemptive or reactive corrective feedback during task are valuable in scaffolding learners’ gradual grammaticalization. This may facilitate learners’ form focus in a broader sense than Skehan’s (1998) emphasis on pre-task form focus and Long’s (1996) exclusive use of focus on form in communication breakdowns.

Though Swan’s (2005) reasoning that L2 exposure-poor learners need to be equipped with sufficient grammar for communicating meaning is logical, Ellis (2009a) argues that they also need more communicative exposure to compensate what their instruction-based environment lacks. To strike a balance, Ellis suggests that meaning-oriented task-based learning can be complemented by rich input, consistent feedback, task design, and focused situational tasks eliciting both learners’ natural language use and specific form. Meanwhile, Ellis cautions that task-based learning may not necessarily suit education settings which uphold knowledge not skill teaching, for example the Confucian-heritage educational philosophy in Asian culture.

Ellis’s proposals and concern subsequently bring forward the effects of task and context. Task and context can be factors influencing learners’ attention to recasts and their potential interlanguage restructuring (Nicholas et al., 2001). Elucidation of the context variable will follow that of the task variable below. The section on the task variable is of considerable length because it is not only a factor affecting learners’ attention to recasts, but also the pivot of eliciting learners’ use of forms for recasts to target.

This section has introduced another kind of developmental readiness. Deriving from Anderson’s (1983) classification of declarative and procedural knowledge in his ACT model, learners’ readiness can be of a relative instead of absolute kind. EFL
learners often have relative readiness for grammar forms, for example they have acquired declarative knowledge of forms but have not proceduralized the use of forms. Swan (2005) argues that EFL learners' unproceduralized use of forms is caused by their deficient grammar foundation. He therefore denies the relevance of task-based learning in the EFL context. Ellis (2009a) however posits that tasks can be meaning and form-oriented at the same time.

3.6 Task

Task is another variable which could cast an effect on the effectiveness of recasts. In general, a principle called transfer appropriate processing posits that it would be easier for learners to retrieve and use the target form in tasks, if the tasks elicit learners' use of the similar cognitive processing as the one they experience in previous focus-on-form learning sessions. Both Spada and Lightbown (2008) and Lyster and Saito (2010) have made use of it to explain findings in their focus on form. Lyster and Saito deduce from their meta-analysis of classroom studies that learners may improve from receiving corrective feedback within the ongoing classroom interaction, if they later use the form in a free oral production task which also demands spontaneous retrieval of the form. Spada and Lightbown also state that the benefit of integrative FFI, which incidentally draw learners' attention to form during communicative activities, would be best shown when learners subsequently retrieve the form also in a communicative task.

From the skill practice perspective, tasks which invite learners' repeated or systematic use or transfer of the form previously targeted by form-focused instructions or corrective feedback in a new context creatively, but under the same content theme, may help foster learners' automaticity (De Ridder et al., 2007).
Automaticity is defined as learners' repeated, stable, but creative use of the form learned before in a new context, without consciously spelling the rules or being aware of the need to use the form from the context (DeKeyser, 2003; De Ridder et al., 2007). Accordingly, tasks adhering to the transfer appropriate processing principle may stimulate learners' automaticized use of the form, and thereby gauge their automaticized usage level.

3.6.1 Differences between the Two Attention and Task Models

Under the discussion of the task demand effect on learners' attention and hence language learning and performance, the single-resource model of limited and thus selective attention (Skehan, 1998; Skehan & Foster, 2001; VanPatten, 1990, 1996) and multiple-resource model of divided attention (Robinson, 2001a, 2003; Wickens, 2007) are the two major conflicting attentional models. Robinson and Gilabert (2007) elucidate the main difference between the two models in terms of their views of the beneficial effects of task complexity on learners' L2 learning and development.

According to Robinson and Gilabert (2007), Skehan's (1998) single-resource model and trade-off hypothesis contend that learners' attentional capacity is globally limited; and any kind of task complexity would unequivocally cause learners' reduced attention to forms. This may then deteriorate either accuracy or complexity as a result of the attentional competition. On the other hand, Wickens's (2007) multiple-resource model and Robinson's (2003) cognition hypothesis conceive attention as consisting of separate resource pools. Tasks being conceptually complex would direct learners' attention to the linguistic resource pool for deploying the form needed to express the relevant conceptual meaning. This may thereby stimulate learners' attempt of the
target form and eventually new form-meaning mapping, leading to greater accuracy and complexity (Robinson, 2005).

The above first comparison may enlighten that Robinson’s attentional model and view of task complexity are more directional than Skehan’s, regarding how a task should be manipulated in order to motivate learners’ use of the target form, and thereby enhance their new form-meaning mapping. This conception is derived from the aforementioned difference between Robinson and Skehan’s attentional theories. Robinson deems that learners’ attention capacity is not limited; it entails different resource pools assuming different cognitive duties to orchestrate learning performance. His view of task manipulation involves increasing a task’s conceptual meaning to direct or stretch available attentional resources to deploy linguistic forms meeting the task’s particular conceptual meaning. Contrastively, Skehan, and also VanPatten (1996), maintain that learners’ attention capacity is limited with meaning overriding form when the two demand attention simultaneously. Thus, Skehan’s view of task manipulation mainly involves pre-task-to-post-task procedural and non-language-based means to nurture additional attentional resources for form, rather than steer existing attentional resources to motivate learners’ use of form during the task.

Robinson’s Triadic Componential Framework (2001a, 2005, 2007; Robinson & Gilabert, 2007) purposely guides task design and control. It includes components of task complexity, task condition and task difficulty. Task complexity is classified as conceptual and procedural demands (e.g. present or past timeframe and increased or decreased planning time), task condition as interactional demands (e.g. one or two-way information flow and sharing content knowledge with interlocutors or not), and task difficulty as ability requirements (e.g. high or low working memory and high or low processing anxiety) (Robinson et al., 2009). Robinson (2001a, 2001b) notes that
the task complexity factor guides task designers along deciding prospectively how a task should appear and work to elicit learners’ certain linguistic behaviours; whereas the latter two factors concern online task design decisions.

Moreover, Robinson’s (2005, 2007) position of attention and perception of task complexity offers developmental opportunities, when directing learners’ attention to the linguistic resource pool and prompting them to use a specific new form required by the task demand. Skehan’s (1998) approach of task manipulation on the other hand does not particularly demand learners to use a specific new or less familiar form in the task. In other words, Skehan’s hypothesis may focus more on controlling task designs to strengthen or proceduralize learners’ access and control of their existing knowledge. Task design according to Skehan’s model may therefore inform learners’ performance and current proficiency; whereas Robinson’s model may facilitate learning through tasks.

Robinson’s cognition hypothesis also accommodates task design components which demand and consolidate learners’ access to and proceduralization of their existing interlanguage (Robinson, 2005; Robinson et al., 2009). Robinson (2005, 2007) on one hand holds that tasks imposing both “conceptual/linguistic” (Robinson & Gilabert, 2007, p.165) and “performative/procedural demands” (p.165) on learners can facilitate learners’ achieving real world communication via prompting both their new and existing linguistic resources respectively. He on the other hand notes that increasing the latter dimension of task demand on learners’ proceduralization of current interlanguage may weaken the facilitative effect of the former on learners’ development of new forms. He therefore suggests, in line with Skehan’s limited capacity hypothesis, that lesser task procedural pressure may free cognitive resources for learners to focus on their linguistic development. The accommodation of both the
developmental and proceduralization views of task complexity may demonstrate the comprehensiveness of Robinson’s cognition hypothesis.

Apart from being more directional, development-oriented and comprehensive, Robinson’s view of task complexity is also considered more favourable to directing learners’ attention to help from input. It is assumed that learners would tend to seek more help from input (Robinson & Gilabert, 2007) when the conceptual demand of tasks increases. Instead of perceiving task complexity as favouring such a learning opportunity (Robinson, 2007) of attending to input, Skehan’s model would predict learners’ attention to be distracted from input when task complexity overloads their attentional capacity.

In Robinson’s (2007) empirical study, Japanese learners playing the role of listeners of their partners’ story-telling exhibited more turns of meaning negotiation via clarification requests and confirmation checks when working on the more complex narrative versions. On the part of the story-telling learners, their increased uptake and use of the pre-modified input containing the target forms in their subsequent narratives were also demonstrated with the more complex task versions. Both the increase in the listeners’ turns of seeking help from negotiation of meaning and the story-tellers’ use of pre-modified input to inform their story meaning in the more complex task illustrated Robinson’s hypothesis that task complexity drives learners to seek more help from input. The pre-modified input provided in Robinson’s (2007) study contained the lexical and syntactic forms relevant to the task linguistic demand, and storytellers could follow the exact forms, partially change the forms, or use their own forms to tell the story to their partners. In the case of following the exact forms, the story-telling learners’ success of using the target forms in the more
complex narrative versions could have been an artifact of the written provision of pre-modified forms and not the benefit of seeking help because of task complexity.

Last but not least, Robinson and Skehan differ in their measurement of speech production. Robinson adopts a more specific measure than Skehan. In addition to employing Skehan’s (1998; Skehan & Foster, 2001) general measure of learners’ speech production according to fluency (e.g. syllables per second), accuracy (e.g. error free C-units) and complexity (e.g. clauses per C-unit), Robinson (2005, 2007; Robinson & Gilabert, 2007) confines the measurement to learners’ use of particular L2 domains expected by certain task conceptual demands. This specific scope of measurement may serve as a more interlanguage-sensitive (Robinson, 2005) and hence relevant indicator of learners’ speech production than gauging their overall error free and complex clauses uttered. For example, Robinson et al. (2009) point out that only specific measure can unveil learners’ extent of using developmentally advanced or non-prototypical tense and aspect in more complex tasks requiring temporal reference. They distinguish simple tasks from complex tasks in the sense that the former entails lower communicative demands and may thereby trigger more use of the pragmatic mode of communication than the morphological mode. In contrast, complex tasks may stimulate learners’ extensive use of grammatical morphology, which is indispensable to task completion and communication success.

General measure which assesses learners’ error-free utterances may stay satisfied with learners’ prototypical use of tense in a complex task but deficient in revealing learners’ development.

Regarding the general measure, Housen and Kuiken (2009) note that the notional and operational definitions of fluency, accuracy and complexity are multidimensional. For instance, complexity can be defined in terms of clause unit,
lexical variety or different verb forms; accuracy in terms of the native standard or community norm; and fluency in terms of pause, false start or number of syllables. These three aspects of speech production can also be measured either by impression ratings or quantifiable indices like frequencies or ratios. Comparatively, the specific measure of learners' certain interlanguage tokens, for example learners' prototypical and non-prototypical attempts of past tense in Robinson's (1995) early study, may entail more clarity as a measuring tool.

3.6.2 Measure of Speech and its L2 and Task Correlates

Besides stating clear about what is being measured, the L2 correlates of complexity, accuracy and fluency need to be made clear for reflecting learners' different commands of L2. Housen and Kuiken (2009) identify that complexity and accuracy signal the current state of learners' interlanguage or internalized L2 knowledge representation, since the former illustrates learners' state of expanding their existing knowledge and the latter reveals their conforming to the target norms. On the other hand, fluency unveils learners' control of their linguistic knowledge. They categorise complexity carrying a L2 correlate as linguistic complexity. Other types of complexity are task complexity, which delineates properties of a language task; and cognitive complexity, which specifies difficulties learners encounter while processing language features, for example input saliency.

Concerning acquisition, Ellis (2009b) elucidates that it can mean acquisition of new linguistic features, restructuring or gaining greater control over existing linguistic features. Moreover, he states that planning may lead to the latter two types of acquisition but not the first type. Ellis however contends that the more complexity learners venture in the task, the more restructuring they may experience and the more
likely they may reach acquisition; whereas *accuracy* may restrict learners from moving forward by controlling and avoiding mistakes. Accordingly, task design may need to be manipulated by means of planning to facilitate learners’ exhibition of complexity and thereby acquisition. Fluency may only show how learners control their existing knowledge, and is hence irrelevant to acquisition.

In addition to specifying the L2 correlates of the three production aspects, Ellis (2009b) also relates them to the effect of task planning on acquisition, for purposes of enforcing appropriate task implementations to facilitate desirable learning outcome. He lists three types of planning: rehearsal, strategic planning and within-task planning. *Rehearsal* is repeating the same task before actually doing it. *Strategic planning* is preparing for the content and language use beforehand without doing the task once. *Within-task planning* is monitoring the content and language use while doing the task either under pressured or unpressured timing. In his review of previous studies on these three types of planning, he concludes that rehearsal fostered fluency and complexity, strategic or *pre-task planning* promoted fluency and facilitated either complexity or accuracy, and within-task planning benefited accuracy and complexity.

Among all task procedural variables, researchers have vastly studied planning and task structure. Earlier studies conjoint with later ones, Ellis (1987) first of all found that learners achieved the greatest accuracy of regular past in the planned writing task, medium accuracy in the planned spoken task, and the least accuracy in the unplanned spoken task. Second of all, in Foster and Skehan’s series of empirical studies on the influence of planning and task structure, Foster and Skehan (1996) illustrated that planning had positive effect on learners’ complexity overall and accuracy when learners had less detailed planning. This multifaceted finding regarding accuracy was due to the study’s methodological contrivance of detailed and
undetailed planning. The more detailed planning condition misled learners to more lexical than formal focus. With the additional variable of task structure, Skehan and Foster (1997) discovered that learners' accuracy was attained in tasks with transparent inherent structure and planning, and tasks with less clear structure led to greater complexity when planning was allowed. In Skehan and Foster's (1999) study, the task condition involving non-simultaneous processing of watching the storyline and then telling it contributed to learners' greater complexity; and greater accuracy in the structured task version under the same non-simultaneous task condition than the unstructured version. Last but not least, precisely pointing to the structure of narrative, Tavakoli and Foster (2008, 2011) and Tavakoli (2009) both found that narratives with tighter structure yielded greater accuracy effect.

As shown above, findings about the pre-task planning effect on complexity and accuracy have been the most unpredictable (Ellis, 2009b; Skehan, 2009). Ellis (2009b) identify variables of learner proficiency and task structure in some studies as possibly moderating the effect of pre-task planning on learners' production, with planning being less effective when given to advanced learners and well-structured tasks. The relevant assistance may have been redundant to advanced learners and well-structured tasks. On the other hand, Ellis (2009b) makes use of Skehan’s (1998) limited capacity model and de Bot (1992) and Kormos’s (2006) L2 application of Levelt’s (1989) L1 immediate speech model to explain the clear trade-off interaction between complexity and accuracy in some other studies.

According to Skehan’s model, when pre-task planning offers rooms for learners to prepare for taking risk in attempting different forms in the subsequent task (complexity), their accuracy of remaining on the safe track to control their conformity to the target norms may be sacrificed due to limited cognitive resources. This
competition between allocating resources to complexity and accuracy is hypothesized to happen in a grammar-encoding component called *formulator* in Levelt's model. Formulator is preceded by a meaning-decoding component called *conceptualizer* and followed by a speech component called *articulator*. de Bot and Kormos state that formulator and articulator undergo automatic operation in L1 immediate speech production, but controlled operation in L2 leading to the conflict between complexity and accuracy. However, Robinson's cognition hypothesis maintains that complexity and accuracy can co-exist through task complexity manipulation.

Judging from the above comparisons between Robinson and Skehan’s attentional and task complexity models, Robinson may seem to outshine Skehan in accommodating the possibility of simultaneously encouraging learners’ form complexity and accuracy under the conceptual complexity task demand. However, Skehan (2009) argues it is the additive effect of different task condition and characteristic control, not task complexity per se as predicted by Robinson’s cognition hypothesis, which leads to the simultaneous increase in accuracy and complexity. Skehan (2009) quotes Tavakoli and Foster’s (2008) finding to support the role of the different separate task manipulations. Basing on Tavakoli and Foster’s study, Skehan claims that a task requiring the integration of both background and foreground information to form a complex storyline and bearing a tight structure with clear beginning and ending would lead to complexity and accuracy respectively. The additive effect of the complex storyline condition and the structured story characteristic may be the underlying reason driving the simultaneous increase in complexity and accuracy. Skehan therefore affirms that task complexity per se, the mere effect of conceptual complexity in other words, may not be a convincing reason motivating the simultaneous increase being discussed.
Skehan (2009) further strengthens his argument by citing Michel et al.’s (2007) study which showed the increase in accuracy but only a marginal change in complexity. Moreover, the increase in accuracy is significant in the monologic task but not the dialogic task. Their partial support of Robinson’s cognition hypothesis concurs with Skehan’s denial of the facilitative effect of conceptual task demand on simultaneously promoting accuracy and complexity. It is noteworthy that Michel et al. also investigated the correlation between conceptual demand in terms of more or less elements in the content and interactional demand in terms of monologic or dialogic. In the case of dialogic interaction, learners’ form accuracy may be raised through triggering more negotiation with interlocutors and reducing learners’ errors through feedback moves. Meanwhile, interaction may mitigate learners’ attempts of complex production by inducing their elliptical answers to interlocutors’ clarification requests and confirmation checks for example. Michel et al.’s study used total number of clauses, subordinate clauses and lexical words to measure complexity; so it can be concluded that the dialogic task demand affected learners’ complex production by triggering their elliptical answers to interlocutors’ responses.

This speculation has been discussed earlier by Robinson’s (2001b) study of a monologic map direction-giving narrative task. It showed that increasing conceptual complexity led to learners’ greater form accuracy and complexity, corroborating Robinson’s cognition hypothesis. Although he only increased task complexity in terms of cognitive demands (planning time, single task, prior knowledge and few elements) and not interactional demands, he discusses that the use of monologic tasks would be favourable to engendering learners’ greater form accuracy and complexity. This echoes Yuan and Ellis’s (2003) claim that dialogic tasks possess interactional variables influencing learners’ performance. Therefore, Michel et al.’s (2007) unequal
results of accuracy and complexity may have been partly due to the monologic and
dialogic variables, not necessarily the conceptual demand variable. Moreover, in
Skehan’s (2009) example of Tavakoli and Foster’s (2008) study, the integration of
background and foreground information in a narrative constituting storyline
complexity is already categorised as a *conceptual demand* (+/- few elements) in
Robinson’s triadic componential framework (Robinson et al., 2009), and the tight
beginning and ending structure of a story constituting well-structuring as a *procedural
demand* (+/- task structure). Skehan’s (2009) claim of the additive effect of separate
task conditions and characteristics can be seen as actually echoing Robinson’s
cognition hypothesis.

3.6.3 Similarity between the Two Attention and Task Models

The convergence of Skehan’s additive effect argument and Robinson’s
cognition hypothesis may subsequently project the intrinsic resemblance between
their attention and task complexity models. Section 3.6.1 has already discussed the
comprehensiveness of Robinson’s model. This section elaborates more on how
Skehan’s model is implicated in Robinson’s model.

Structurally, the multiple-resource model comprises six different resource
pools dividing learners’ attention along three main task dimensions (Robinson, 2003;
Wickens, 2007): 1) the *spatial versus verbal resource pools* for regulating codes of
processing; 2) the *auditory versus visual resource pools* for regulating modalities of
processing; and 3) the *manual versus vocal resource pools* for regulating responses of
processing. Robinson (2003) affirms that resource competition only occurs within
rather than between the different resource pools. Unlike the single-resource model
which assumes learners’ holistically limited attentional capacity and overloaded
attention under task complexity, the multiple-resource model considers that learners’
cognitive capacity will not be overloaded when task complexity disperses learners’
cognitive load across their separate resource pools.

As mentioned before, Robinson (2001a, 2003; Robinson & Gilabert, 2007)
possits that the more complex the task demand is regarding the language requirement,
the more likely learners’ cognitive resources will be pushed and directed to a specific
resource pool regulating linguistic encoding. This seems to be in the same vein as
Swain’s (1985, 2005) output hypothesis, which predicts that learners will be pushed
to test or process their developing linguistic forms under the demand of output
production in fulfilling communicative needs. With attention divided and directed to
different specific resource pools, Robinson assumes that learners will be able to attend
to linguistic codes required for task completion.

Thereby, this resource-directing dimension of task complexity, in Robinson’s
(2001a, 2003; Robinson & Gilabert, 2007) term, can be interpreted as favouring
learning opportunities regardless of task complexity. Taking an example, tasks
eliciting the use of past tense to construct time meaning, which is considered as more
complex than the use of present tense due to the reduced contextual support of here-
and-now or displacement of time and space (Robinson, 2001a, 2005), may push or
direct learners’ attentional resources to their linguistic coding pool. Learners are then
predicted to eventually adopt target-like forms obligated by the communicative task.
In contrast, Skehan and Foster (2001) perceive coding complexity as one of the task
complexity features constituting learners’ cognitive overloading.

On the other hand, Robinson (2001a, 2003) takes into account possibility that
certain task complexity variables disperse learners’ attentional resources in a non-
specific manner, and induce them to disable specific linguistic coding. For example,
tasks with less planning time and hence more complex than otherwise will probably weaken learners’ formal performance (Robinson, 2001a). This coincides with Skehan and Foster’s (2001) analysis of the task complexity effect on attention. They suggest that time pressure, for instance, will deplete learners’ limited attention and lead them to engage predominantly in meaning decoding (Skehan, 1998; VanPatten, 1996). This resource-depleting or dispersing dimension (Robinson, 2001a, 2003; Robinson & Gilabert, 2007) informs that intensifying task complexity can also jeopardize learners’ formal learning.

3.6.4 Task and Recasts

Studies of task-based learning (Ellis, 2003, 2005b; Robinson, 2001a, 2005, 2007; Skehan, 1998; Skehan & Foster, 2001) advise that different degrees of task complexity help learners gradually restructure their interlanguage through switching their focus to form in task performance. Though the different forms learners exhibited in completing the task may not be target-like, the formal output elicitation involved encourages learners to test or process the different formal possibilities based on their internal syllabus (Swain, 1985, 1995, 2005). To effectively circumvent non-target-like form, formal feedback intervention has been suggested useful in destabilizing learners’ interlanguage to attain native-like proficiency (Long & Robinson, 1998; Long, 2007; Swan, 2005). Moreover, Schmidt (2001) propounds that it is learners’ noticing of the gap between their errors and the target forms, facilitated by the juxtaposition of errors and models in recasts and subsequently learners’ cognitive comparison (Doughty, 2001; Long, 2007), trigger their interlanguage destabilization and then restructuring.
Along this line of research on attention to formal input, Robinson and colleague (Robinson, 2001a, 2005, 2007; Robinson & Gilabert, 2007) propose the advantage of increasing task complexity in facilitating interlocutors' delivery as well as learners' uptake of focus on form techniques. They reckon that the more complex the task is along the resource-directing dimension, the more motivated learners will be to seek help from input, and more easily will they attend to communicative-redundant forms in the task (Robinson, 2005). On the other hand, Skehan’s (1998; Skehan & Foster, 2001) information processing approach holds that task complexity should somehow be restricted to channel learners’ attention to form in input.

Gilabert (2007) empirically illustrated that learners’ attention to form, measured in their monitoring and self-repairing grammatical and lexical forms, was enhanced as task complexity built up. He also comparatively revealed that narrative outshined other tasks with less clear structure and higher demand of devising content from scratch (instruction and decision-making tasks) in guiding learners’ attention to form. The number of repairs was observed higher in the complex versions of all the three tasks (narrative, instruction and decision-making tasks). This was consistent with Robinson’s cognition hypothesis that task complexity motivates more attention to form; and reflecting Wicken’s (2007) multiple-resource attentional model that learners can attend to both form and task meaning. However, Gilabert et al. (2009) found that task complexity has significant effect on triggering non-recast interactional techniques in the narrative and instruction tasks. They attributed the low occurrence of recasts in the study to their EFL learners’ lack of training on recasting, due to their grammar-oriented and less communicative language education. Their educational culture also did not prepare them to recast partners’ errors.
Apart from the above studies drawing relationships among task, attention to form and recasts, Revesz (2009) pioneers a very recent empirical study to explore learners' formal performance facilitated largely by recasts as task complexity grows. Although Robinson’s (2003, 2005, 2007; Robinson & Gilabert, 2007) resource-directing task variable suggests that task complexity can direct learners' attention to deploying their verbal linguistic resources to distinguish L1 and L2 forms for attaining L2 accuracy (Robinson, 2005), Robinson himself mentions and Revesz (2009) later queries that the competition between learners’ L1 and L2 resources within the verbal linguistic pool may deplete the limited capacity of that single pool. Revesz raises the potential resource-depleting problem, faced by L2 learners, when increasing task complexity along the resource-directing dimension. To be precise, she wonders if L2 learners' verbal linguistic resource pool will always have enough capacity for accommodating the resource-directing task complexity.

Synthesizing the different views about resource competition during L2 task production, Skehan (1998) maintains that human's attentional resources are non-differentially limited, and learners' fluency, accuracy and complexity all compete for scarce resources. VanPatten (1996) states that form and meaning compete for limited resources under their simultaneous processing. Robinson (2003) argues that attentional resources are only limited within the different resource pools as proposed in the multiple-resource model. Robinson makes use of the interference theory to consolidate his hypothesis that resources are limited within each resource pool. The competition between L1 and L2 within learners' verbal resource pool is one of the instances constituting learners' within-pool limitation when performing L2 tasks:

From the perspective of interference theory, explanations linking relative ease or difficulty of L2 comprehension, or different characteristics of L2 production, to task
demands may be more legitimately framed in terms of confusion and cross-talk between codes (of L1, interlanguage, and L2 syntax, morphology, semantics, and phonology / orthography) within specific resource pools during task performance, rather than in terms of global capacity limitations. (Robinson, 2003, p.646)

The cross-talk theory between L1 and L2 codes would apply to the case of EFL students. Their limited exposure to the target language may make their L1 resource code appear more frequent and salient than that of the target language, and therefore wins over the L2 code for more attentional resources (Schmidt, 1990, 2001) during task production. Apart from the competition within the verbal pool, another within-pool competition may come from the vocal resource pool.

The vocal resource pool structurally allocates resources for responses of auditory perception, and the manual resource pool separately for responses of visual perception (Robinson, 2003; Wickens, 2007). In a case which the single vocal pool has to allocate resources for learners’ oral responses to both their auditory perception of recast correction (e.g. incorporating the use of target form into following utterances) and their visual perception of cartoon-strip picture meaning (e.g. speaking about the pictures) simultaneously, the resource competition within the vocal resource pool may be incurred. Revesz’s (2009) findings exactly illustrated that the spoken narrative task with the presence of visual aid actually affected learners’ learning from recasts.

Revesz’s (2009) study, a recent one which investigates the specific relationship among task complexity, recasts and L2 learning, may lend empirical support to the rationale of controlling the effect of task on learners’ learning from recasts. The key findings were that learners gained more from recasts when the treatment did not provide photos as visual aid to their oral description than when
recasts were with photos; and that non-recast learners benefited more from the presence of photos than otherwise in using past progressive. Revesz uses Robinson’s cognition hypothesis (2001a, 2007), task complexity variables (2001a, 2005, 2007), and proposition of the link between task complexity and the efficacy of recasts (Robinson & Gilabert, 2007) jointly to justify these two findings.

Revesz (2009) employs Robinson and colleague’s two contradictory resource-depleting task effects to support her finding of the efficacy of recasts. Robinson posits that the increase of task complexity along the resource-depleting dimension, for example the dual task demands of learners’ speech production and their memorization of meaning previously presented in pictures in Revesz’s study, would raise learners’ cognitive burden. Learners would then seek more help from recasts to compensate the resource-depleting demand; and Revesz’s finding that learners with no photo support benefited more from recasts demonstrated this advantage of the resource-depleting task complexity. Moreover, according to Wickens’s multiple-resource model, the absence of pictures may avoid the resource competition between learners’ vocal responses to both their visual (pictures) and auditory perception (recasts).

Meanwhile, Revesz mentions Robinson’s other argument that the resource-depleting task demand, realised in the dual task demands in her study, may distract learners’ attention from recasts. This was because of the task switching and scheduling operations incurred when learners had no visual support to inform task meaning. Revesz uses this disadvantage of the resource-depleting task demand to advocate the more significant contribution of recasts in facilitating the non-picture group’s gains shown in their posttests than the contribution of pictures in lessening learners’ resource-depleting burden for the picture-recast group.
Revesz's finding discussed above showed the overriding effectiveness of recasts over the hypothesized effectiveness of providing picture support to advise task meaning. However, her further finding that the picture and non-recast group outperformed the non-picture and non-recast group in using the form may substantiate the benefit of decreasing the resource-depleting task complexity via picture support, and freeing attentional resources for learners to meet the resource-directing task complexity of using L2 past progressive. This finding may guide the use of picture prompts to control the resource-depleting task effect, to allow learners enough attention to the target form and recasts.

This section has mentioned the transfer appropriate processing principle and compared the difference between Robinson (2003, 2005, 2007; Robinson & Gilabert, 2007) and Skehan’s (1998) attentional and task complexity models, to inform the proper design of tasks. The former may be more helpful in guiding the design and control of tasks. It may also be more likely to motivate learners’ use of complex linguistic forms and seeking help from recasts. Robinson’s adoption of the specific measure of learners’ use of form may more closely measure learners’ interlanguage development than the general measure that Skehan usually adopts. The facilitative use of picture prompts and recasts in tasks was shown in Revesz’s (2009) study.

3.7 Modality

Even with the different control measures in tasks, the modality of tasks may still affect learners’ attention to learning. The modality of speaking, as opposed to writing, imposes greater demand on L2 learners and even greater on EFL learners because of their receptive-based classroom instruction and written exam-oriented learning. Ellis and Yuan (2005) identify that the extent to which L2 learners possess
procedural knowledge of a form in rapid online speaking determines whether the 
modality of speaking is a variable to learners’ task performance. They further classify 
that both L1 children and L2 immigrant learners tend to have higher procedural 
knowledge of target language in speaking over writing due to their limited literacy 
level; but EFL learners are likely to perform better in target language writing than 
speaking because of more time to deploy their controlled processing in writing.

Wolff (2000) strongly argues that writing is more powerful than speaking as a 
language learning tool because of their different cognitive demands. Firstly, writing 
triggers deeper language processing than speaking because it requires more mental 
efforts from learners in producing precise language; secondly, writing entails no 
online time constraint when compared to speaking, so writers can spend more time on 
deploying controlled processing in pondering their language use; thirdly, writing is 
more likely to raise language awareness than speaking because of its neatly visualized 
segments rather than fading aural signals. All these suggest the different input and 
output effects of writing and speaking modalities on learners’ language learning and 
use.

3.7.1 Effect of Modality on Learning of Form

    Regarding learners’ language learning from input, VanPatten’s (1990) study 
found that simultaneous processing of meaning and form imposed difficulty on 
learners, with learners’ attention to form of lower communicative value affecting their 
attention to meaning at the same time. Therefore, VanPatten (1994, 1996) proposes 
the input processing approach to make certain formal input salient enough for 
learners to notice, in order to conduct form-meaning mapping. Wong (2001) adds the 
investigation of modality, aural as well as written modes, to VanPatten’s (1990)
original study. She asked different learners to listen for the content only, a lexical
word plus content, and a grammar word plus content; in addition, she asked learners
to do all these in a reading/written mode. The results showed that learners did not
have difficulties with simultaneously attending to both form and meaning in the
written mode; while learners had the same difficulties with the aural mode as
VanPatten’s (1990) learners. Wong’s (2001) study empirically showed that learners’
performance in the written mode was better than that in the aural mode.

Leow et al.’s (2008) more methodologically cautious study also found the
non-interfering effect of attention to different lexical and linguistic forms on attention
to meaning in the written input mode. They included a control group attending to
meaning only, a group attending to the monosyllabic lexical item *sun* in Spanish, a
group attending to the meaningful and grammatical item *it* in Spanish, a group
attending to the grammatical item *the* in Spanish, and a group attending to the
morphological form –*n* in Spanish. The use of the monosyllabic item *sun* was to
equalize the syllabic saliency of all the forms, a refinement to VanPatten’s (1990) use
of the polysyllabic word *inflation* in Spanish. The pronoun *it* in Spanish, entailing
both meaning and grammatical functions, was also used as a refinement to
VanPatten’s (1990) study to probe closely the form-meaning connection made by
learners. Leow et al. attributed the non-significant effect of attention to forms varied
in their meaningful values on meaning comprehension to the possibility that some
learners backtracked the non-fading meaning of the written text. After sidetracking
their attention to form momentarily to circle the forms assigned while reading the
text, learners may have turned back to the reading and answered the subsequent
comprehension questions. Moreover, the think-aloud protocol looking into learners’
concurrent attention to meaning and form during reading showed their low depth of
processing the form. This may also have led to the non-interfering effect of attention to formal input on learners’ simultaneous attention to meaning input in a written text.

Apart from the effect of modality on learners’ learning from input, modality also has an effect on learners’ learning from output. Swain (1985, 1995, 2005) and Swain and Lapkin (1995) contend that output can push learners to notice the gap between what they can and what they are expected, test their language hypotheses in the need of output production, and gear interlocutors’ feedback to their specific language output. Niu’s (2009) study explores the effects of producing written and spoken output on learners’ attention to form. One group of learners engaged in oral reconstruction of a passage they read collaboratively with their partners, and another group of learners did the passage reconstruction in the written form collaboratively with their partners. The collaborative reconstruction process involved pairs of learners discussing different aspects of language—content, grammar, structure and vocabulary. The results showed that learners in the written output task attended more to language forms than learners in the oral output task.

In sum, Wong’s (2001) study examines the modality effect on learners’ attention to formal input; and Niu’s (2009) study investigates the modality effect on learners’ attention to form in their discussion and reconstruction output. Both suggest that the speaking mode has lesser facilitative effect on learners’ attention to form. However, there have so far been few studies probing the relation between modality and corrective feedback (CF). Sheen (2010), one of the few, compares the effects of oral and written CF on L2 learners’ learning of the English indefinite article a and the definite article the. The effect of oral CF was performed through giving recasts and oral metalinguistic feedback; and that of written CF was through written direct correction and written metalinguistic feedback. Since recasts and written direct
correction provided learners with the correct form directly, both are equivalent except the difference in modality. More obviously, oral and written metalinguistic feedbacks are equivalent except their modality. Sheen’s definition of metalinguistic feedback is different from that of Lyster (2004). Lyster conceives metalinguistic feedback as a type of prompt which withholds the correct form and elicits learners’ own reformulation; whereas Sheen turns metalinguistic feedback into one that gives the correct form plus a grammar explanation. The results showed that all the feedbacks, except recasts, led to learners’ improvement in their L2 English article accuracy.

Sheen (2010) attributed the results to the explicit-implicit distinction rather than modality, because of the non-significant difference between the effects of the oral and written metalinguistic feedbacks. These two kinds of metalinguistic feedback only differ in modality. However recasts and written direct correction differ not only in modality, but also explicitness due to the clearly explicit visual correction provided by the latter. All oral feedbacks were given immediately after learners’ errors in front of other classmates; whereas all written feedbacks were returned in a delayed manner after teachers collected and corrected their individual written output. This online-offline distinction is a prominent difference between the spoken and written modes. Because the written mode gives time for learners to respond and process, Sheen originally hypothesized that all the written feedbacks would excel their spoken counterparts in influencing learners’ L2. Contrastively, modality was not the driving force of the relevant influence. Recasts entail the advantage of being implicit and non-interruptive to the ongoing communication according to Long (1996, 2007). Without any explicit transformation, recasts are in and of themselves implicit. With oral and written metalinguistic feedbacks exerting the same effect and written direct correction
being facilitative, the ineffectiveness of recasts is better explained by their implicitness than spoken mode.

Although Sheen's (2010) study adds value to the field by illustrating the clear mediating effect of explicitness rather than modality, her study committed certain biasing problems disfavouring recasts. First of all, as investigated by other researchers, recasts can be turned explicit by certain characteristics, for example reduced recasts and higher-pitched reformulation on the errors. By controlling the level of explicitness of the four feedbacks in the study, the factor of modality could have been examined more independently. Second of all, the test instruments used - speeded dictation test, writing test, and error correction test - were all in written form and testing explicit knowledge. Recasts, which were implicit in spoken form and not as explicit as other feedbacks, may have exerted more effectiveness if the tests were oral spontaneous output testing learners’ implicit knowledge.

3.7.2 Effect of Modality on Task Planning

Studies on how pre-task or strategic planning and within task or online planning (Ellis, 2005b) facilitate learners’ task performance (fluency, accuracy and complexity) have largely been conducted in the speaking mode (Foster & Skehan, 1996; Ortega, 1995, 1999, 2005; Sangarun, 2005; Skehan & Foster, 1997, 1999; Wendel, 1997; Yuan & Ellis, 2003). More planning time given to foreign language learners to implement speech production would be especially advantageous, because they possess limited procedural knowledge of vocabulary and grammar (Ellis & Yuan, 2005). Hence, more careful planning during speaking to create time for the ongoing grammar formulation, as in the writing process, would ease learners'
performance constraints. Speaking however allows less time for learners to access their linguistic resources via controlled processing (Ellis & Yuan, 2005).

The effects of the different modalities of speaking and writing on task performance can be explained by the two representative models of speaking (Levelt, 1989) and writing (Kellog, 1996). Levelt’s (1989) *speech production model* hypothesizes that L1 speakers undergo 1) conceptualization of identifying communication goals and organization of the preverbal structure of the communication goals, 2) formulation of internal speech in the shape of grammar and vocabulary, 3) articulation of internal speech, and 4) monitoring of articulated speech. De Bot (1992) distinguishes Levelt’s L1 model from L2 speakers’ case by arguing that L1 speakers operate their formulation and articulation processes almost without attention; but L2 speakers, especially those with limited proficiency, often need to exert their working memory to execute the formulation and articulation stages via controlled processing. Therefore, to avoid pressuring L2 learners’ speech processing, Ellis and Yuan (2005) suggest the use of careful within-task planning to permit time for learners’ controlled and monitoring processing.

Kellog’s (1996) model of writing involves similar processes to those of the speaking model, which are formulation, translation, execution and monitoring; except speaking exerts more real-time pressure on learners’ working memory than writing (Ellis & Yuan, 2005). In other words, writing entails some time-out from online production for learners to monitor pre and post-execution output; and its visual output makes mistakes more salient for monitoring to amend. As Ellis and Yuan (2005) found that their Chinese learners achieved more complexity (grammar and lexis) and accuracy in writing than oral production, careful online planning, a privilege that writing modality grants, would be advantageous for learners to advance and monitor
their speech. About the relationship between planning and speech measurement, Yuan and Ellis (2003) earlier found that pre-task planning specifically contributed to learners’ greater grammatical and lexical complexity, and online planning promoted their greater grammatical complexity and accuracy. This may also reflect that pre-task planning bears a more focused effect in targeting complexity.

As shown in Ortega’s (2005) post-task interview findings of her learners’ actual use of pre-task planning in her previous studies (Ortega, 1995, 1999), low-intermediate learners reported that they made use of retrieval strategies during their pre-task planning to retrieve existing form and solve lexical and verbal morphological problems for later performance. However, some of her learners also spoke about the difficulty to transfer pre-task planning on lexical and grammatical form to online performance. Ortega (1999) earlier reasoned that this could have been because the online task demand overrode the pre-task planning effect. Moreover, same as Wendel’s (1997) finding that learners did not actually use the pre-task planning time to focus on form, Ortega’s (2005) learners articulated that they focused more on meaning because of the worry of not expressing comprehensible enough meaning to their listeners. Despite learners’ different foci during their pre-task planning, Ortega (2005) concludes that pre-task planning is beneficial to learners by offering them time and opportunity to self-regulate their linguistic resources before the upcoming speech task.

Wendel (1997) and Ortega (2005) thus show that learners may not focus on linguistic form during their pre-task planning; in contrast, Sangarun (2005) explores whether specific instruction guiding learners’ focus during pre-task planning would foster their use of attention as intended. The study investigates the different effects of pre-task planning with different guided instructions on meaning focus, form focus and
meaning-and-form focuses, and no pre-task planning on learners’ speech task performance. Although the three experimental conditions outperformed the control group, only learners under the meaning-and-form instruction condition did the best. Sangarun’s (2005) finding is directional in terms of advising the benefit of guiding learners’ use of the pre-task planning time via giving both meaning and form instructions.

This section has explained how the different modalities of tasks and feedbacks affect learners’ form learning and benefit from task planning. VanPatten (1990), Wong (2001), Leow et al. (2008) and Niu’s (2009) findings suggested that learners were able to simultaneously attend to form and meaning when the task was in the written mode. The online cognitive burden caused by the spoken mode may lead to learners’ weakened attention to meaning and form. Task planning may therefore be an option to facilitate learners’ attention to form. Yuan and Ellis (2003) advise the functions of the different kinds of task planning. Oretga (1999, 2005) however found that learners could not transfer their attention to form in pre-task planning to the actual speaking task, due to the online cognitive burden caused.

3.8 Context / Setting

The physical constraints of context and setting may hinder the facilitative effect of different task designs and control measures on learners’ attention to form or learning from corrective feedback. Two recent meta-analysis studies provide an overview of the effect of these variables. Same as Lyster and Saito (2010), Li’s (2010) meta-analysis also synthesizes the effect of corrective feedback and the different factors contributing or mediating the effect across different past L2 studies. The difference is that the former concludes that instructional context does not
contribute to the effectiveness of corrective feedback; whereas the latter attributes the
different effect sizes of corrective feedback and the different variables to the
overarching influence of instructional context and research setting. Li (2010)
interprets *context* as about foreign language or second language context, and *setting* as
about laboratory-based or classroom-based setting. Other studies used the two terms
in the opposite way. The different interpretations however do not contribute any
significance to the following discussion.

Li has taken a cautious stance when reporting the findings by considering
other possible factors behind the findings. First of all, the effect of corrective
feedback was found exerting medium effect size and maintaining over time, which is
of a lower effect size than previous meta-analyses (e.g. Russell & Spada, 2006;
Mackey & Goo, 2007). However, Li alerts readers that the different inclusion and
exclusion criteria of choosing past studies to meta-analyse may have led to the
different effect sizes. Second of all, explicit correction had larger effect size than
recasts and metalinguistic feedback across studies. However, Li attributed this to the
studies’ use of mechanical drills as tasks to test the effect of explicit correction. The
same discrete-item format in both the drilling task and explicit correction treatment
may have favoured the effect of explicit correction. Li’s reasoning concurs with the
transfer appropriate processing principle and Norris and Ortega’s (2000) meta-
analysis finding. On the other hand, recasts had a larger long-term effect than its
short-term effect. Li related this to the long-term effect of implicit feedback. Explicit
feedback was found more effective on immediate and short-delayed posttests in
developing explicit knowledge, and implicit feedback worked better on long-delayed
posttests in developing implicit knowledge. Again, Li cautions that different
methodology and learner characteristics may have caused the different effects.
Notwithstanding Li’s skepticism about the unsure findings above, she reports the finding that the foreign language context contributed larger effect than the second language context. She explains that foreign language learners are keener on grammar instruction and error correction; whereas second language learners incline more to developing communicative skills to survive in their L2-dominant community. Another finding is that the laboratory-based research setting carried a substantially larger effect than the classroom-based setting. Li reasons that learners often received feedback in laboratories on a one-on-one and discrete-item basis intensively throughout the treatment session and carefully with extraneous variables tightly controlled. Thereby, constrained mechanic drills were found mostly in laboratory studies and free communicative activities in classroom studies. The intensiveness and controlled setting of laboratories may have also enhanced the salience of feedback. Because free production in communicative activities may not constrain or obligate learners’ use of the target form, learners may have avoided it to secure themselves from errors. Therefore, Li suggests that constrained construction tasks may accurately gauge learners’ use of the form via providing an obligatory context. For other effects, the larger effect of short treatment over longer treatment and the larger effect of L2 English over other types of L2 in the past corrective feedback studies were attributed to the setting and context respectively. Short treatments were often conducted in laboratories and L2 English was often taught in intensive language programmes. Overall, setting and context may have been the fundamental factors building the effect of corrective feedback.

Among specific studies about the effect of context, Sheen (2004) launched a descriptive yet comprehensive one. She investigated the different effects of different instructional classroom contexts on teachers’ use of corrective feedback, learners’
corresponding uptake and repair, and the relationship between the feedback types and learners’ uptake and repair. Through comparing observational data obtained from Lyster and Ranta’s (1997) French immersion classrooms in Canada, Panova and Lyster’s (2002) ESL classroom in Canada, Ellis et al.’s (2001) intensive ESL classrooms in New Zealand and her EFL classrooms in Korea, Sheen found that teachers in all the four meaning-based instructional contexts likewise used recasts most out of other feedbacks identified in Lyster and Ranta’s (1997) taxonomy.

However, with regard to the differences, teachers in the New Zealand ESL and Korea EFL classrooms used more recasts than the Canadian classrooms, with recasts appeared in the Korea EFL classrooms most frequently. Sheen (2004) explains that the most frequent pattern of recasts in the EFL context was derived from its free conversation-based curriculum; and the nature of recasts of not interrupting ongoing communication was precisely shown compatible with the context. Concerning learners’ rates of uptake and repair, the New Zealand ESL and Korea EFL learners responded with more uptake and repair than the Canadian classroom learners. Sheen (2004) attributed this finding to learners’ higher proficiency and education background in the former two contexts. This reasoning coincides with Mackey and Philp (1998) and Philp’s (2003) studies, as discussed in the section on learners’ level, that level can be a factor in influencing learners’ attention to recasts.

Specifically about learners’ uptake and repair of recasts, learners in the New Zealand ESL and Korea EFL contexts exhibited more uptake and repair following recasts than the other two contexts. This is justified by the fact that the two Canadian classrooms used recasts for negotiation of meaning, offering less opportunity for learners to respond before topic continuation within the ongoing meaning discourse. The other two classrooms nevertheless employed more explicit recasts—one or two
target features, reduced reformulation, intonational emphasis, and form-focused recasts with opportunity for uptake. The explicit implementation of recasts may offset the ambiguity of recasts in being indistinguishable from non-corrective repetitions occurred in meaning discourse (Lyster, 1998a). In Doughty’s (2001; Doughty & Varela, 1998) view, rather than discrediting recasts as ambiguous, the techniques of delivering recasts should be adjusted in a more salient way, to draw learners’ contingent errors to their short-lived cognitive window, for immediate cognitive comparison between their errors and the target forms.

The additive salience in recasts recommended by Doughty (2001) and its effectiveness in motivating learners’ greater uptake and repair in the more form-focused contexts Sheen (2004) studied may also be explained by Lyster and Mori’s (2006) counterbalance hypothesis. Lyster and Mori (2006) propose the counterbalance hypothesis to account for their findings that the Japanese immersion learners exhibited uptake and repair of recasts more than prompts but the French immersion learners showed the opposite; and predict which classroom context is suitable for recasts or prompts. Instead of investigating the differential effectiveness of recasts and prompts within the same context (Lyster, 2004; Ammar, 2008; Ammar & Spada, 2006), Lyster and Mori (2006) inject the context variable to their comparative study of recasts and prompts. They reason that the Japanese immersion learners benefited more from recasts because of their teachers’ more structured and accuracy-focused way of implementing communicative teaching. This form-focused way of communicative teaching seems incongruent with the unplanned and contingent meaning-focused nature of recasts. Furthermore, Lyster and Mori explain that the French immersion learners gained more from form-eliciting prompts because of their teachers’ contrastive approach of meaning-based and non-interruptive way of
communicative teaching. With each interactional feedback type counterbalancing the instructional context it appeared in, Lyster and Mori argue that learners from either context were then able to benefit from the outstanding and thereby unambiguous or salient feature of recasts and prompts.

Sheen (2004) and Lyster and Mori's (2006) comparative studies may comprehensively inform that SL/FL classroom contexts with grammar-oriented teaching and proficient learners would offer more optimal opportunity for learners to notice and benefit from recasts than their meaning-oriented counterparts. This conclusion seems to share the same early view as Nicholas et al. (2001). They suggest that the facilitative effectiveness of recasts hinges on classrooms’ form-focused orientation; and EFL classrooms, which incline to formal grammar teaching, may most likely benefit learners with form-focused recasts. However, a counterintuitive finding may have emerged in their studies, which is for example meaning-oriented classrooms were not found actually meaning-oriented and form-oriented classrooms were not shown truly form-oriented. Thereby, it may become difficult to predict the effectiveness of recasts accurately based on the overarching characteristic of a classroom. Both Ellis et al. (2001) and Loewen’s (2004) observational studies on ESL meaning-oriented classrooms in New Zealand and Tsang’s (2004) examination of the Hong Kong EFL classrooms may thoroughly illustrate the counter-intuition.

As shown in section 2.3 of chapter 2, Ellis et al. (2001) and Loewen (2004) explore the classroom pattern of uptake and the characteristics influencing learners’ uptake and successful uptake in the same ESL adult classroom setting in New Zealand. The overall rates of uptake as well as success of uptake in both studies were higher than Lyster and Ranta’s (1997) widely cited French immersion classroom study with children. Another difference from Lyster and Ranta’s study is that the New
Zealand studies also probe into pre-emptive focus on form techniques which are initiated by a teacher or student to address a linguistic point explicitly, apart from reactive ones under which a teacher responds to a student’s linguistic errors (Ellis et al., 2001). Being pre-emptive or reactive was one of the characteristics classifying the different focus on form techniques employed in the New Zealand classrooms and predicting learners’ uptake and success. Other characteristics established in the studies were complexity, language focus, timing, directness, and source of the form focus episodes. The authors attributed their higher rates to their form-oriented population of East Asian learners and pre-emptive form-oriented approach, despite the overarching meaning-oriented ESL school context.

Although form-focused classroom elements were shown to have driven learners towards focusing more on form, Tsang (2004) did not find such predictable advantage of form-focused elements in an EFL context for learners in Hong Kong to gain from recasts. In accord with Lyster and Ranta (1997) and Lyster’s (1998b) descriptive studies on meaning-focused immersion classrooms, Tsang (2004) also observed that recasts were most frequently used but least effective in leading to learners’ uptake and successful repair among other negotiation-of-form feedbacks (prompts). The Hong Kong secondary EFL classrooms Tsang studied carried a mixture of meaning and form focuses. This implies that neither meaning nor form classroom orientation was a facilitator of learners’ gain from recasts, not even when the classrooms carried conventional EFL form-focused pedagogical and learning practice.

Lyster and Mori’s (2006) counterbalance hypothesis may then fail to explain Tsang’s (2004) finding that EFL form-focused orientation did not help direct learners’ attention to repair their errors signaled by recasts. However, this conflict could be
because the accuracy-focused Japanese immersion learners in Lyster and Mori’s study were already used to their classroom practice of oral repetition after teachers’ instruction. Hence, they were more trained to repeat after recasts, though repetition after recasts is considered as a redundant discourse move within meaning-based communication (Lyster & Ranta, 1997).

Lack of training on responding to recasts is perhaps a more precise reason than the overarching classroom orientation behind learners’ failure in responding to recasts in the form of uptake and repair. Therefore, with no guarantee or confidence that learners actually know or are familiar with the response manner to recasts, signs of uptake and repair may not be used to indicate learners’ noticing and learning of form. To reiterate, Lyster and Ranta (1997) do not hold firm that the use of immediate uptake of recasts offers definitive measurement of learners’ learning of the form. They only suggest that, based on Swain’s (1985, 1995, 2005) claim that output is facilitative in pushing learners to move from semantic to syntactic processing, producing uptake may help learners practise using the form for later automatized retrieval. However, this suggestion is rather speculative. Especially with Asian EFL learners or even Asian language native speakers who tend to be reticent in their culture, overt uptake of feedback or provision of feedback during ongoing conversation may be a challenge. Therefore, relatively reliable evidence from the pre- and-posttest measurement may refine the flaw of using uptake to measure reticent interlocutors’ attention to and learning from feedback during interaction.

In Iwashita’s (2003) quasi-experimental study investigating the effectiveness of positive (model) and negative (recasts) evidence given by Japanese native speakers to Japanese foreign language learners on two Japanese structures, a pretest, immediate posttest and one-week-later delayed posttest were employed to gauge the effect of
feedback during the communicative task treatments. Recasts were shown more facilitative to learners’ short-term gain than other conversational moves in the posttest scores, though positive evidence dominated in terms of frequency. Iwashita’s study also demonstrated that nearly half of learners’ non-targetlike utterances were ignored during interaction. It was explained that Japanese native speakers’ culture of viewing silence as a virtue and avoiding confrontation could have been the reasons; thus only less than half of the non-targetlike utterances were either negotiated or recast. This indicates that Asians’ reticent culture may conceal observable responses during interaction.

The variable of context, involving meaning or form orientation, amount of relevant training on focus-on-form techniques, and learners’ learning or communicative culture, has been shown to pose certain impact on learners’ attention to and use of feedback. Oliver and Mackey (2003) articulate the importance of taking diverse contextual factors into account when analyzing the different dynamics of classroom interaction; for example learners’ non-targetlike initial turns, opportunities for teachers’ feedback, and learners’ actual use of feedback. In their observational study, the categorical context of interactional classroom has been dissected into different components—content-focused, communication-focused, classroom management-focused and explicit language-focused contexts. They found that communication-focused episodes, which appeared more unstructured and unplanned along the natural flow of communication, had led to most learners’ non-targetlike initial turns. This was attributed to the unplanned nature of the communicative flow and therefore more cognitively demanding for learners to cope with. Furthermore, the explicit language-focused episodes triggered teachers’ feedback provision and learners’ use of feedback 85% of the time, and were often given via recasts. The
results showed that more form-oriented contexts may ensure teachers' feedback and 
learners' use of feedback to a certain extent, due to their being more structured and 
hence more salient and less cognitively demanding when compared to communicative 
contexts.

Although Oliver and Mackey (2003) exhaustively examine the different types 
of classroom interactional settings, they view their results cautiously by reminding 
that the umbrella variable of context can consist of many other subtypes, for example 
"interlocutor effects, pedagogical settings, social factors, and experimental contexts" 
(p. 521). They also admit that laboratory settings can give rise to a more consistent 
interactional environment than classrooms. This seems to echo Nicholas et al.'s 
(2001) earlier claim that the laboratory setting can shape recasts as more useful input 
to learners than the classroom setting.

Nicholas et al. (2001) contribute a thorough review of L1 and L2 past studies 
on recasts to the second language acquisition field. It reconciles the dispute over the 
different findings showing the variable effectiveness of recasts. Nicholas et al. 
conclude that findings revealing the effectiveness of recasts need to be taken 
cautiously, because different factors may have caused the enhancing or mitigating 
effect to the success of recasts. The different factors producing the different 
effectiveness of recasts, as discussed in the last and current chapters, are context, 
learners’ level or developmental readiness, the implementation of recasts, 
measurement of the effectiveness of recasts, target form and task.

For context, communicative classroom interaction which triggers high 
frequency of recasts may lead to learners’ difficulty with distinguishing corrective 
recasts from non-corrective repetitions, especially when both are usually accompanied 
by conversational acknowledgment. Experimental settings may however contrive the
operation of recasts in more consistent, intensive and restrictive (single or fixed set of
target features) manners. Moreover, the form or meaning orientation in a classroom
(ESL or EFL) or laboratory setting may lead learners’ attention to either form or
meaning accordingly. Learners with higher proficiency or more ready development
may excel in benefiting from recasts. For implementation, recasts can be made more
explicit with intonational emphasis on the error or target form to attract more
attention. The different measurements of learners’ attention to and use of recasts, for
example immediate surface uptake versus posttest performance, can indicate the
multifaceted effectiveness of recasts. The effectiveness of recasts also varies along
language aspects such as phonological, morphosyntactic and lexical forms. Last but
not least, task, which has been given quite extensive discussion above, is also
identified by Nicholas et al. (2001) as potentially determining if meaning is first of all
clear enough for learners to free attention to recasts.

This section has been suggesting that context affects learners’ attention to
form. However, Gass et al.’s (2005, 2011) comparative study of the classroom and
laboratory settings revealed that there was no difference between the two in
influencing learners’ focus on form. They employed three typical interaction
techniques to analyse the contextual variable: 1) negotiation of meaning in terms of
using confirmation checks, clarification requests and comprehension checks, 2)
language-related episodes (LREs), and 3) recasts. The results illustrated that the
foreign language classroom setting and the laboratory setting did not cause any
significant difference to learners’ amount of using negotiation of meaning, focus-on-
form episodes and recasts in their interaction with each other. Although setting did
not constitute an effect, task was found significantly affecting learners’ attention to
form. Learners interacted with each other in dyads to accomplish three information-
exchange activities—picture-difference, map, and consensus. Classroom learners engaged in the tasks during regular class time with teachers walking around as in a normal instruction session; whereas laboratory learners did the tasks in a separate session without the presence of the teachers. The study yielded the significant results that both the picture-difference and map tasks led to learners’ increased use of negotiation of meaning, LREs and recasts than the consensus task.

The findings have to nevertheless be viewed with caution because recasts were particularly low in number, learners were foreign language (FL) learners sharing the same L1, and the laboratory learners were left alone interacting by themselves. Gass et al. acknowledge that if learners in the study were second language (SL) learners, they would have negotiated for meaning more with each other in classroom because of their different L1. Moreover, if the laboratory learners interacted with the teachers instead of their peers, instances of negotiation and other means of focus on form would have been delivered more consistently. If these were the cases in the study, then the two settings may have differed in their effects. As the authors conclude, the setting variable itself entails considerable complexities.

As Gass et al. (2005, 2011) point out, there is a wide range of different foreign language classrooms. Philp and Tognini (2009) state in their review the different variables involved in conducting interactions in a FL classroom. Because of the restricted target language exposure in a classroom and limited classroom time that FL teachers have, teachers often minimize students’ opportunities to interact with each other or with them in L2. For example, teachers often do not expand their interaction with learners to offer them negotiation for meaning opportunities. Moreover, teachers usually provide more scaffolding to assist learners with expressing meaning because of the limited time and their limited proficiency. Regarding attention to form, FL
teachers often give more corrective feedback than SL teachers because of their form-focused instructional curriculum. Philp and Tognini found in their review of past studies that prompts, recasts and explicit correction were the most common feedbacks used by FL teachers. In learner-learner interaction, FL learners tend to prefer using L1 because they share the same L1 and possess limited L2 proficiency. To facilitate L2 learning through tasks, the authors suggest FL learners receive vocabulary or formulaic language scaffolding, familiar tasks, as well as tasks requiring use of certain form to achieve certain communicative meaning.

This section discusses the effect of context on learners’ benefit from recasts. Both Sheen (2004) and Lyster and Mori (2006) demonstrated that learners in form-oriented classrooms learned from recasts more than meaning-oriented ones. Lyster and Mori hypothesized that the meaning-based nature of recasts may have appeared more outstanding and hence effective in form-oriented contexts. However, the reason could have been that Lyster and Mori’s learners had training about responding to recasts. Iwashita (2003) noted earlier that the culture of learners may affect their response to recasts. Gass et al. (2005) however found no interaction between the laboratory and classroom contexts, because of the ways the two contexts had been made. Studies like Ellis et al. (2001), Loewen (2004) and Tsang (2004) also showed that their claimed-to-be meaning and form-oriented classrooms had not actually been what was originally claimed to be.

3.9 Conclusion

This chapter has drawn the scope of discussion more narrowly in order to examine specifically the different ways of delivering and researching the effectiveness of recasts in the EFL context. The use of measurement tests is advised to
be knowledge-appropriate. In other words, controlled tests obligating discrete rules and forms are best measuring learners’ explicit knowledge or the effectiveness of explicit feedback targeting learners’ memory of discrete rules and forms. Free oral production tests are on the other hand suitable for measuring learners’ implicit knowledge or the effectiveness of implicit feedback targeting learners’ spontaneous feel for the form used to communicate the corresponding meaning. Moreover, the use of feedback or recasts is suggested to be optimal as input and output practice through offering intensive exposure and performance-based error correction, especially in the context of EFL. The use of perception-indicating tools has been shown bearing value in probing the internal features of recasts effecting learners’ learning, though their certain imperfections need to be taken cautiously. Considering the deficient level of EFL learners, the use of grammar teaching and communicative tasks is encouraged to be held hand in hand to help EFL learners attain sufficient formal foundation for communicative use in fulfilling real-world demands. The use of linguistic-complex but well-controlled tasks, pre-task planning for speaking tasks, and more form-oriented contexts may better prepare EFL learners with benefiting from recasts. These may compensate EFL learners’ lack of relevant exposure or training.

3.10 Research Gaps and Questions

Advancing from the above variables as represented in the literature review in chapters 2 and 3, the present study attempts to contribute to the dearth of research to date by empirically exploring the comparative effects of recasts with different levels of saliency. Both Nicholas et al. (2001) and Sheen (2006) have previously identified this research gap; so the present study ventures to test the fundamental hypothesis that the more salient recasts are, the more effective they are to promote L2 development.
In other words, comparing the implicit and explicit types of recasts becomes the study’s focus. The different implementations of recasts in different studies, according to their different degrees of explicitness and implicitness (Braidi, 2002; Doughty & Varela, 1998; Doughty & Williams, 1998; Han, 2002; Loewen & Philp, 2006; Long, 2007; Long & Robinson, 1998; Lyster, 1998a, 1998b; Lyster & Ranta, 1997; Mackey & Philp, 1998; Nicholas et al., 2001; Sheen, 2006), demonstrated the different effects of recasts. Yet, with the exception of Nassaji’s (2007, 2009) studies, no clear finding has emerged to date about the comparative effects of explicit and implicit recasts within the same context.

Furthermore, investigating the comparative effects of explicit and implicit recasts may isolate the different variables besides recasts which also contribute to the effectiveness involved. In sections 2.6 and 2.7 of chapter 2, the variables of frequency and saliency were shown to have made certain type of feedback more explicit and hence more effective than the other. This may thereby obscure the effect of the feedback without these variables, as in the difficulty of determining whether it was the saliency variable or recast which contributed to the effectiveness of Doughty and Varela’s (1998) corrective recast. However, very few studies have examined the different effects of feedbacks with and without reinforcing variables.

Another gap that the present study aims to fill is the under-researched context of Hong Kong EFL secondary school. Tsang’s (2004) study has been introduced as one which examines recasts in the Hong Kong context. However, it lacks a comparative angle in viewing recasts. Although Hong Kong students have been taught English since their kindergarten years, their extensive use of it outside classroom is very limited. Based on section 2.8.2 of chapter 2, they can be considered having rather sophisticated explicit knowledge of English grammar rules but deficient
implicit knowledge of English usage in communication. However, there have not been sufficient studies looking into Hong Kong learners’ relative readiness of L2 knowledge.

One example of illustrating Hong Kong learners’ relative readiness of L2 knowledge is their use of past tense. As discussed in section 3.5 of chapter 3, EFL learners are often taught through fixed instructional syllabi assuming their linear language development and lack the exposure of using the target language for communicative purposes in tasks. Accordingly, Hong Kong learners become relatively familiar with the rules of past tense but relatively unfamiliar with using past tense. Regarding the two dimensions of past tense, section 2.8.2 of chapter 2 has identified regular past as less perceptually salient and acquired later than irregular past because of its phonological ambiguity in speech stream and rule-bearing feature.

About the interaction between the two past tense dimensions and feedbacks, Yang and Lyster (2010) employ Skehan’s (1998) dual-mode system to conclude that recasts are more suitably used to target irregular past and prompts for regular past. With the less explicit nature of recasts in not demanding learners’ uptake, they may not be effective in eliciting learners’ uptake of the computational rule processing-based and phonologically ambiguous regular past. On the other hand, irregular past, which is lexical-based and learners’ quick retrieval will most likely be triggered, may be easily noticed in recasts. Regardless of all these individual findings, very few studies have linked them together and probed the use of recasts to help Hong Kong learners’ use of irregular past in communicative tasks. As Ellis (2009a) noted, balancing formal and communicative focuses is crucial for EFL learners; hence applying consistent recasts to assist Hong Kong learners with attending to form in
communicative tasks may be advisable. This complementary approach has also been under-researched in Hong Kong.

Prompts have been clearly demonstrated their greater effectiveness than recasts in eliciting learners’ attentive immediate responses; and recasts often fail to do so, leading to either learners’ silence or mimicking. As noted by researchers quoted in sections 2.3 and 2.5 of chapter 2, uptake to recasts may not be a viable indicator of learners’ noticing and learning; the pre-and-posttest measurement of learners’ learning may be a more reliable choice. Investigation of Hong Kong learners’ noticing and learning from recasts across time may therefore be more worth doing, especially when Hong Kong learners lack the relevant training of responding to recasts. Very few studies have considered this limitation of Hong Kong learners.

Another limitation of Hong Kong learners that very few studies have thought of is their familiarity with explicit tests eliciting memory of discrete item rules, but unfamiliarity with implicit tests eliciting their feel for the form instead of memory of the form in free production in their education. Therefore, use of communicative tasks implicitly eliciting Hong Kong learners’ use of the target form may need to be used in studies, to probe more accurately their communicative use of the form.

Based on these research gaps identified, corresponding research questions are formed as follows:

1) Are unenhanced consistent recasts facilitative in promoting Hong Kong learners’ implicit knowledge of L2 form usage in communicative task in the short run?

2) Are unenhanced consistent recasts facilitative in promoting Hong Kong learners’ implicit knowledge of L2 form usage in communicative task in the long run?

3) Are enhanced consistent recasts facilitative in promoting Hong Kong learners’ implicit knowledge of L2 form usage in communicative task in the short run?
4) Are enhanced consistent recasts facilitative in promoting Hong Kong learners’ implicit knowledge of L2 form usage in communicative task in the long run?

5) Are enhanced consistent recasts more facilitative than unenhanced consistent recasts in promoting Hong Kong learners’ implicit knowledge of L2 form usage in communicative task in the short run?

6) Are enhanced consistent recasts more facilitative than unenhanced consistent recasts in promoting Hong Kong learners’ implicit knowledge of L2 form usage in communicative task in the long run?

3.11 Conclusion of Chapters 2 & 3

It seems that the absolute effectiveness of any kind of corrective feedback, be it recasts or prompts, cannot be warranted because of the presence of different co-existing variables: complexity of communicative context, improper measurement, deficiency of controlled setting, frequency, perceptual salience, proficiency level or developmental readiness, target form, test, task and modality. They may have contributed to causing or engineering certain deceptive effects. These variables have been discussed heightening, obscuring or misrepresenting the effectiveness of recasts and prompts. Notwithstanding their interfering impacts, the aforementioned variables may serve as useful indicators to advise the proper application of recasts to maximize L2 learning. The extensive section on task above specifies some useful indicators developed by researchers studying task design and control, to inform the proper application of task as instrument to measure the effectiveness of recasts.

As Ellis and Sheen (2006) review, recasts can take many different forms and functions, not necessarily corrective as normally expected. This is due to 1) their both communicative and didactic roles depending on their implementation in terms of
being implicit or explicit, intensive or not, and their different characteristics; 2) their context of operation in terms of being utilized in meaning or form-oriented classrooms or laboratories, and the frequency and consistency of recasts; and 3) learners’ level or developmental readiness in terms of being ready to treat language as objects (morphosyntactic focus) or tools of communication (lexical or phonological focus). Concerning measurement of recasts, the effectiveness of recasts has been shown throughout the entire literature review as leading to subsequent learning and ultimate acquisition from learners’ across-time performance measurement, not learners’ fleeting immediate response to recasts. Unless learners’ uptake can be shown equal to learners’ subsequent and lasting use of the form in focus, as in Loewen’s (2005) study which found that learners’ posttest results reflected their accurate recall of previous forms in their uptake, then signals of uptake can be treated as indicating learners’ acquisition.

Chapters 2 and 3 have reviewed and identified the interfering variables as well as useful guidelines on how the effectiveness of recasts can be affected. The following chapter addresses how these variables of recasts were controlled to facilitate learners’ attention to the target form and recasts in the present study. The inherent effectiveness of recasts may thereby be evaluated more fairly.
Chapter 4 Methodology

4.1 Introduction

This chapter gives details of both the quantitative and qualitative research design and methodology employed in addressing the research questions outlined in chapter 3. It identifies the different variables to be inserted into the research questions, to guide a well-controlled study for the relevant investigations.

The present study made use of an experimental design to investigate Hong Kong EFL learners’ progress in using irregular past tense in their spoken narratives, from pre to posttests after receiving the different types of recasts. Moreover, stimulated recalls were used to explore learners’ cognitive experiences at the time of receiving recasts and content-only feedback. The study context is set in a Hong Kong secondary school with a group of 15-year-old students in their third year.

Responding to or receiving recasts or immediate feedback in spoken discourse may be alien to Hong Kong secondary school students, because of their lack of spoken practice and limited training on responding to recasts. Because of learners’ novelty with responding to recasts and the repetition-like response to recasts, the present study adopted the pre-and-posttest measurement of the effectiveness of recasts. Moreover, as with Iwashita’s (2003) Japanese participants, the present study’s Chinese learners bear the culture of reticence. The present study was hence motivated to use a more controlled experimental setting to elicit Chinese learners’ considerable amount of output. The present study however referred to learners’ immediate responses to recasts when analysing their perception towards the use of past tense, their mistakes, and recasts during the feedback sessions.

The following sections will elucidate 1) the theoretical and empirical justification of implementing both the experimental and introspective designs in the
present study; 2) the context of the present study; 3) the participants’ characteristics and their selection criteria and methods in both the experimental and introspective designs; 4) the study’s choice of target form; 5) the study’s use of instrument for both the experimental and introspective designs; 6) the study’s data collection procedures and logistics in both the experimental and introspective designs; and 7) the methods adopted for the study’s quantitative and qualitative data analyses. All these may provide insights into the complete study design. The pilot study conducted prior to the actual study will also be delineated, and how the latter has been refined from avoiding limitations found in the former. The research hypotheses will be formulated at the end of the chapter.

4.2 Theoretical and Empirical Justification of the Research Design

4.2.1 Quantitative Research Design

The initial philosophy underpinning the adoption of an experimental approach is Lyster and Mori’s (2006) counterbalance hypothesis. The implementation of a kind of feedback whose principle is contrastive to the pedagogical orientation of a context or setting is hypothesized to exert effectiveness. Accordingly, recasts, which are typically communicative and thereby implicit, are suggested to be effective in drawing the attention of learners in a form-oriented context or setting to form, because of their unusual presence. Moreover, Yoshida (2008) acknowledges the pedagogically and socially favourable characteristics of recasts, which make recasts suitable to the limited class time and deficient students constraints of an EFL context. However, as Iwashita’s (2003) study indicates earlier, learners’ lack of training or exposure to recasts in their culture of education may compromise their gain from recasts. To harmonize both views, an experimental approach controlling the amount
of exposure to recasts may be an option. It can make use of quantity to compensate form-oriented learners’ novelty with recasts, while the use of recasts in their form-oriented context or setting is retained.

The rationale behind choosing an experimental design, not an exploratory one, is the guarantee of the former in ensuring sufficient amount of feedbacks to be given, with learners’ output and error instances to be elicited. As suggested by studies in section 3.8 of chapter 3, experimental settings excel classroom settings in delivering consistent, intensive and salient feedbacks, with the unpredictability in natural classrooms controlled to the minimum. Moreover, learners’ output and their use of the target form can be expected to a certain degree, by controlling the type of tasks or tests as the motivation. Sections 2.7 and 3.5 of the last two chapters also advise that consistent feedbacks or formal input may scaffold EFL learners with deploying forms in fulfilling communicative task demands, which they are deficient in. Experimental settings may favour such a condition where learners need to be guided in producing semi-spontaneous output elicited by certain tasks. Classroom settings however rely heavily on the spontaneous course. Their difference in spontaneity leads to the question of whether experimental findings can apply to real-world classrooms, which will be addressed in chapter 8.

Inspired by the above reasoning, the present study chose the experimental approach as its quantitative dimension. According to Porte’s (2002, p.64) definition, the present study’s research design can be classified as a kind of quasi-experimental design, because it involved both experimental and control groups. The present study contained two experimental groups receiving feedback on past tense in different degrees of explicitness (normal recast and corrective recast), and one control group receiving feedback on the story content only. Normal recasts are defined as the
classical or unenhanced type of recasts employed by Long (1996, 2007); and corrective recasts as the refined or enhanced type of recasts used by Doughty and Varela (1998). They were compared to distinctively view the effect of the enhanced variable. The present study deliberately arranged the control group to receive feedback on the story content instead of no feedback at all, to avoid the participants being singled out too much from the two experimental groups. Moreover, participants had not been randomly selected or assigned to take part in the study (Porte, 2002, p.64). A pre-test was conducted to ascertain that the non-randomized participants all started at the similar level, to possibly reduce the history effect of participants. More detail on the participants and the pre-test administered will be provided in sections 4.5 and 4.7 respectively.

Based on Schmidt’s (2001) noticing hypothesis and researchers’ (Sheen, 2006; Egi, 2007a; Kim & Han, 2007) characterization of attention-getting recasts, the present study took the essential relationship between noticing, learning and implementing recasts explicitly to support its fundamental hypothesis. The present study predicted that isolated, declarative, single-error-focused, reduced, substituting (Sheen, 2006), and enhanced recasts are more facilitative to EFL Hong Kong Chinese students’ learning of irregular past tense than implicit recasts. The implicit recasts in the present study also had the above features, except the enhanced features of repeating learners’ errors with intonational emphasis and subsequently placing intonational emphasis on the target form. This is derived from Doughty and Varela’s (1998) implementation of corrective recasts. Different from Sheen’s (2006) conception, the explicit recasts used in the present study carried the combined use of another feedback type—preceding repetition. Injecting these extra doses of explicitness to recasts is hypothesized to be more effective than implicit recasts.
The deliberate effort of giving explicit recasts to EFL Chinese students may be necessary, because their attention to the use of form in recasts may need to be drawn and enhanced before they can learn the form effectively. Attention as the requisite of learning (Schmidt, 1990, 2001) has bearing on the EFL Hong Kong Chinese students’ case. First of all, EFL Hong Kong Chinese students have limited practice and exposure to the target language under their EFL education, which relies heavily on restrictive form-oriented classroom instructions (Carless, 2007). Second of all, past tense in English is an alien grammar feature in students’ tenseless L1 (Yang & Huang, 2004). According to the single-resource attentional model or information processing model (Skehan, 1998; Skehan & Foster, 2001; VanPatten, 1996), Chinese learners may easily switch to past time lexical adverbials to replace the use of past tense morphology. Either model explains that this is to avoid the low communicative value of past tense morphology causing heavy cognitive load to learners’ limited attention, when dealing with both form and meaning during a task.

In line with Schmidt’s (2001) position that awareness is crucial when learning new L2 form, aiding the Hong Kong Chinese learners’ attention to past tense usage may be necessary because past tense is an unfamiliar form in their L1. Moreover, the use of explicit recasts may enhance learners’ attention to the surface feature of past tense through providing target models. Corresponding to Schmidt’s (1990, 2001) view of verbal report as revealing learners’ noticing, learners’ uptake of recasts was used to indicate their noticing the surface target form (Lyster & Ranta, 1997; Sheen, 2004). Stimulated recall may verify the link between uptake and noticing.

Under the broad identification of quasi-experimental design, the involvement of different variables in the present study contributed to further design classifications. They are namely between-group design, under which the present study examined
different participants assigned to three different independent feedback groups—
*normal recast, corrective recast, and control group (independent variables)*; and
*repeated measures design*, under which the present study obtained the three different
independent feedback groups’ past tense performance from their spoken narratives at
different time intervals—*immediate post-test and delayed post-test (dependent
variables)*.

However, the research design of the present study would have appeared
potentially confounding if the design classifications were settled with the above
description. In other words, further elaboration regarding the group and time variables
may help illuminate how potential extraneous variables were restricted from
distorting the results. For example, the control group, which only received feedback
on the story content rather than the use of past tense, was included in the present study
to help ensure its validity.

The introduction of a control group was to check if the two experimental
groups’ increased or decreased use of past tense in their spoken narratives was really
because of the two feedbacks (normal recast and corrective recast); or merely because
of other factors. For example, the practice effect of doing similar narrative tasks
across time, the history effect of experiencing other school engagements during the
data collection period, the maturation effect of developing the skill of using past tense
across time regardless of the feedback, or the instrumentation effect of the narrative
tasks in giving too obvious hints to participants in using past tense (Porte, 2002, p.78).

More detail on the task or test instrument will be given in section 4.7; the exact
happenings during the data collection period will be reported in section 4.8.

The present study also applied the different time intervals (immediate post-test
and delayed post-test) to help sustain the study’s validity. A delayed posttest was used
to illustrate the inherent effectiveness of the two feedbacks on past tense (normal recast and corrective recast) over time. Without it, the effectiveness of the two feedbacks measured by the immediate posttest may only reflect their provisional effectiveness shortly after receiving the treatments. The inclusion of a delayed post-test may therefore secure the validity of the effectiveness results. With all the above considerations, the present study chose to adopt a 3 x 2 factorial research design, as visualized in table 1 below.

Table 1—3 x 2 Factorial Research Design

<table>
<thead>
<tr>
<th>Time Intervals</th>
<th>Feedback Types</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal recast (on past tense)</td>
</tr>
<tr>
<td>Immediate post-test</td>
<td>Corrective recast (on past tense)</td>
</tr>
<tr>
<td>(spoken narrative)</td>
<td>Control (on story content)</td>
</tr>
<tr>
<td>Delayed post-test</td>
<td>Past tense usage ratio in spoken narratives (past tense usage relative to each participant’s different length of narrative)</td>
</tr>
<tr>
<td>(spoken narrative)</td>
<td></td>
</tr>
</tbody>
</table>

This 3 x 2 factorial design allowed the present study to cover two observable effects on participants’ use of past tense in their spoken narratives at the same time. They were the feedback or group effect and the time effect. The different feedback treatments on past tense (normal recast and corrective recast) have clearly been defined as the independent variables from the very start, and predicted to exert different corrective effects on participants’ use of past tense in their spoken narratives. The different post-treatment checkpoints (immediate post-test and delayed post-test) of gauging participants’ lasting positive or negative progress in using past tense in their spoken narratives served however as both dependent and independent variables. The principal dependent variable in the study was participants’ past tense usage ratio, which is the number of their use of past tense tokens relative to the different lengths of their spoken narratives.
The two time points (immediate post-test and delayed post-test) would primarily be the dependent variables, since participants' performance of using past tense in their spoken narratives at different time was assumed to be dependent on their feedback treatments. However, one of the weaknesses of introducing different time points would be the possibility that participants' increased or decreased use of past tense was due to their maturation or frustration across time, instead of the feedback treatments. A factorial research design will hence accommodate concurrent investigations into the two effects (Porte, 2002, p.83) of feedback and time on participants' use of past tense. Their interaction effects were tested by a specific statistical tool which will be delineated in section 4.9.

The above theoretical explanations of an appropriate experimental research design applicable to the present study may substantiate the present study's quantitative approach of its hypotheses. The quantitative experimental approach was adopted to assist the present study with gaining an "objective, generalizable, ... comparable" (Hauser, 2005, p.306), measurable and hence rather unambiguous picture of the participants' use of past tense across different time points after receiving different feedback types.

Moving from specifying the framework of the study's quantitative approach, some past studies using the similar quantitative approach will be discussed to provide a critical evaluation of the research design and reference to the present study. As discussed in chapter 2, the weakness of using learners' immediate uptake and repair in classrooms to examine their learning from recasts is derived from their uncertainty of reflecting learners’ noticing. The use of individualized tailor-made tests to probe learners’ learning from their exact uptakes of the previous incidental feedback has also been critiqued as testing learners’ memory instead of generalized use of the form.
Therefore, methodologically, the present study chose the use of pretest-posttest measures in experimental research to conclusively gauge learners’ development of extensively using the form.

Both Doughty and Varela (1998) and Han’s (2002) studies on recasts adopted the pretest-posttest measure and controlled implementation of recasts in terms of their consistency and intensiveness. However, the former’s use of recasts was accompanied by classroom content instructions within the same classroom context, under which learners’ performance in posttests may have been gains from recasts, classroom instructions or both. To eliminate the intervening variable of classroom instructions, the present study’s experimental approach did not include any instructional elements; recasts were the sole intervention in an outside-classroom context. However, the regular classes that learners in the present study were having at that time may have exerted some intervening effects. This is when having a control group becomes helpful to disentangle the effects of recasts and classroom instructions. Doughty and Varela’s study also had a control group; but the combined context with recasts given in the same ongoing class may have obscured the separate effects of recasts and instructions. Although the classroom instructions were content-based, learners may have learned some use of the target form implicitly through the teachers’ instruction delivery.

Both Doughty and Varela (1998) and Han’s (2002) studies included a control group. However, the mere reception of their control groups was regular classes, with no participation in any activities during the treatment sessions of the experimental groups. This may have contributed to recasts being compared to nothing; and made the control group learners feel different from the experimental group. The isolation felt by the control group may particularly happen when all participants are from the
same school and same form communicating with each other about the study. Bearing this in mind, the present study assigned the control group to also receive feedback, but on their narrative content only during the recast treatment sessions of the other two experimental groups.

Speaking of using narratives, the present study employed Han’s (2002) same approach of using narratives prompted by cartoon strips to elicit learners’ use of tense and thereby the use of recasts to target their tense errors. However, the use of written narratives in addition to the oral mode as test instruments in the study was not imitated by the present study. As advised by the transfer appropriate processing principle in section 3.6 of chapter 3, tasks which are similar to the previous form-learning tasks in terms of triggering learners’ same cognitive processing may more likely elicit learners’ use of the target form. With oral recasts only given to learners’ oral narratives and none to their written narratives, the present study doubted the relevance of Han’s use of written narratives to gauge learners’ learning from recasts. Since oral recasts were given to learners’ oral narratives in the present study, only oral narratives were adopted as posttests.

Furthermore, Han’s use of written narratives as pre-task preparation for learners’ subsequent oral narratives may have attenuated the spontaneity of learners’ use of tense to convey the same narrative meaning. In other words, learners may have used their memory to re-construct the same narrative meaning in the oral mode. Such a practice would have defeated the present study’s purpose of probing learners’ feel for the use of tense to mesh with the temporal meaning of a historical narrative. Therefore, the present study only included oral narratives eliciting learners’ spontaneous use of past tense, though they were prepared with the narrative meaning via reading a Chinese summary of the narrative before. Han’s practice also seems
contradictory to her target of looking into learners’ procedural level of using tense in real operating conditions (Johnson, 1996), rather than their declarative level of simply retrieving tense rules which they were developmentally ready for. Unlike Han’s concurrent use of random recasts targeting forms other than tense as distracters, the present study kept the focus of all recasts on past tense, because the consistency effect of feedback has been claimed by Ellis (2009a) as facilitative to EFL learners.

Doughty and Varela (1998) and Han’s (2002) experimental studies only investigated one form of recasts. Nassaji’s (2009) study of inspecting the different forms of recasts and elicitations in terms of explicitness may on the other hand lend closer reference to the present study’s focus on enhanced and unenhanced recasts. These studies have already been discussed in chapter 2; what is worth discussing here is how the present study can learn from their methodologies. Apart from the shared features of implementing the explicit and implicit form of the same feedback under the same context and conducting the study outside classroom and class time individually, the present study learned from some methodological doubts in Nassji’s study and thereby underwent different administrations. First of all, Nassaji’s study adopted individualized tailor-made tests to gauge learners’ learning of feedbacks given in interaction targeting the same incidental errors they committed in the pre-interaction test. However, the present study was not of an incidental nature and its target form was pre-selected; so the posttests in the present study did not have to be individualized.

Second of all, Nassaji’s testing of learners’ recognition and correction of the same errors committed in the same output in both the immediate and delayed posttests may incline more to testing their memory than any fruitful learning. Another aspect of relying on learners’ memory was the use of a written description in the pretest as
preparation for the subsequent oral description in the interaction treatment stage. Learners were even encouraged to orally describe the picture scenarios close to what they had written. This reinforcement given to learners’ description may have caused their heightened awareness of errors; and learners may have thereby avoided them as much as possible, giving less room for feedback delivery. The present study doubted the sufficiency of feedbacks in Nassaji’s study.

Third of all, Nassaji only used the different feedbacks to target learners’ same errors that reoccurred in the treatment session from the previous pretest. This was to eliminate the possibility that the errors were learners’ one-off slip of the tongue of their known forms. Errors which represent learners’ real interlanguage gap were hypothesized to keep occurring and therefore need to be restructured. Learners’ fruitful learning from the feedbacks may thereby be facilitated. However, the present study did not take this into consideration because of the uncertainty in what forms learners would generate in their different spontaneous narrative output every time. This uncertainty and the sole dealing of the same exact errors would cast difficulty on delivering sufficient amount of feedback. Moreover, examining learners’ learned but unproceduralized use of form was precisely the present study’s focus.

Inevitably, the experimental approach involves some limitations threatening its validity. This section has indicated that a control group and a delayed posttest were included in the study to ensure the validity that learners’ progress was dependent on the feedback they received. Section 4.8 will particularize other measures taken within the data collection procedures to minimize harms possibly done to the experimental design’s validity. Section 4.9 will reveal how the present study endeavored to attain reliability in its experimental design via the transcription and coding systems developed.
4.2.2 Qualitative Research Design

Another way that the present study used was a qualitative approach which looked into learners' cognitive perceptions of the different feedback types they received during the treatment sessions. The initial philosophy underpinning the use of an introspective approach is the findings of the different L2 classroom studies discussed in section 2.3 of chapter 2 about the limitation of recasts. They attributed the ineffectiveness of recasts to their failure in triggering learners' clear noticing of the form because of their ready-made reformulation for learners, when compared to other response-eliciting feedbacks. Communicatively, repetition of interlocutors' reformulation could be learners' inattentive mimicking (Gass, 2003); and learners' silence could be their attentive response but a way to avoid the inappropriateness to interrupt the ongoing communication (Braidi, 2002). Cognitively, immediate uptake of recasts with the exact reformulation given just seconds before may not require much processing from learners (Trofimovich, Ammar and Gatbonton, 2007).

Methodologically, Mackey (2006) articulates the difficulty of using learners' immediate responses to recasts as surface indicators to reveal learners' covert noticing of recasts and learning of forms. Leeman (2007) and Egi (2010) therefore encourage the combined use of learners' perception and subsequent tests to probe learners' noticing and learning. Examining learners' perception of recasts is essential because different qualities of immediate responses to recasts, even inaccurate ones, could commonly indicate learners' noticing (Egi, 2010).

*Stimulated recall,* which involves interviewing procedures guiding learners' introspection into their own thought processes during a prior activity (Gass & Mackey, 2000, p.1; Lyle, 2003), was chosen to operate the qualitative approach of the present study. It was used to obtain insights into learners' awareness of the feedback
given to their past tense or narrative content, as well as their other cognitive activities at the moment of receiving the feedback and narrating the story.

An example of the conventional stimulated recall interview procedure would be:

Student/Learner (recast session video excerpt): There are ten suns.  
Teacher/Researcher (recast session video excerpt): There were ten suns.  
Student/Learner (recast session video excerpt): There were ten suns.  
Teacher/Researcher (stimulated recall session in LI): What were you thinking at that time?  
Student/Learner (stimulated recall session in LI): Nothing, I was thinking of the story meaning at that time.

The introspective method of stimulated recall was used to supplement the quantitative approach’s inadequate coverage of learners’ individual cognitive processes at a specific moment (Hauser, 2005). Although the experimental design controlled extraneous variables as many as possible and provided objective evidence of learners’ development of past tense, it may have only illustrated an overall picture of their development. In other words, the experimental investigation may not have been in-depth enough in demonstrating what was inside learners’ heads at the time of receiving feedback. Learners’ cognitive encounters during the time of feedback may serve to unveil underlying causes of learners’ post-treatment performance changes. However, it has not been confirmed that learners can bring to consciousness their noticing state. The following will characterize the general purpose and operation of stimulated recall in second language research, its strengths and weaknesses; and evaluate some past empirical studies utilizing stimulated recall, as theoretical and empirical justification of adopting stimulated recall in the present study.

First of all, the need to obtain learners’ naturalistic account of their own thought processes during the actual moments they received feedback may justify the use of stimulated recall in the present study. Apart from outshining immediate recall
as discussed in section 3.4 of chapter 3, stimulated recall is also better than other qualitative measures. Since stimulated recall is not as intrusive as process-tracing or online recall methods for example think-aloud protocols in potentially affecting learners’ ongoing task (Gass & Mackey, 2000, p.18; Lyle, 2003), it appears compatible with the present study’s purpose of maintaining near-naturalistic investigation into learners’ cognitive processes. Moreover, because of its non-intrusiveness to online tasks which learners’ cognitive processes are based on, stimulated recall may not require training learners how to simultaneously verbalize what they think during the task (Gass & Mackey, 2000, p.18). Learners’ recalling their thought processes immediately after tasks also makes stimulated recall appear more effective than delayed post-task interview, which probably induces memory decay by interviewing participants not immediately afterwards (Gass & Mackey, 2000, p.18). In light of the study’s purpose of obtaining learners’ vivid account of their cognitive encounters during the feedback time and the suitability of stimulated recall, the present study chose to adopt stimulated recall as its research method leading to qualitative data analysis.

As detailed by Gass and Mackey (2000), stimulated recall involves the use of tangible stimuli in presenting recent past events visually or aurally to trigger the researcher’s question prompt and the learner’s recall of a cognitive activity at a specific moment. The present study’s use of stimulated recall resembles the procedures implemented by Mackey, Gass and McDonough (2000) most. In their study, stimulated recall followed experimental treatment sessions given to learners’ spoken output of the task; and it was also employed to complement the quantitative view of feedback perceived by learners in the task. Further step-by-step account of the present study’s stimulated recall procedures will be provided in section 4.8.
From the earlier paragraph on stimulated recall's advantages over process-tracing methods and delayed post-task interviews, the strengths of stimulated recall can be described as non-intrusive to learners' ongoing tasks and thereby maintaining learners' naturalistic experience and account (Gass & Mackey, 2000, p.18; Lyle, 2003) of the actual context. It also avoids memory decay when being implemented as immediately after the task as possible (Gass & Mackey, 2000, p.18; Lyle, 2003).

Moreover, stimulated recall allows specific probing into learners' specific inner or tacit thoughts during specific moments of a task (Lyle, 2003), which may counter the potentiality that quantitative findings obscure learners' tacit experience or knowledge at the time of treatment. Though capturing learners' thoughts of specific moments of a task, stimulated recall usually gets access to learners' tacit cognitive processes via indirect probe, rather than directly asking learners to "comment on or generate their own perceptions of their cognitive processes" (Lyle, 2003, p.872). These advantages may justify the present study's choice of using stimulated recall.

Apart from its credited strengths, stimulated recall entails weaknesses which need to be resolved by researchers in order to secure research reliability and validity. The weaknesses are mainly: 1) researchers' question prompts may trigger learners' new reflections instead of pure recalls of the recent past, 2) learners' memory decay if recalls are conducted in later time, 3) learners' bias of presenting themselves most favourably to suit researchers' purpose or researchers' bias of guiding or structuring learners to recall what is preferred to be the answer, 4) stimulated recall findings may not be generalized to a larger population, and 5) learners' different ability of verbalizing their recalls (Gass & Mackey, 2000; Lyle, 2003).

Attempts considered in the present study to overcome limitations 1)—3) and 5) above were avoiding question prompts being too specific to trigger learners' new
responses, hint to learners what is expected, or guide learners’ responses; implementing stimulated recall interviews as soon as possible after the task; using students’ first language to make them feel more comfortable with verbalization; and allowing students’ any responses. However, limitation 4) may be an unavoidable one since the stimulated recall method aims at exploring learners’ individual cognitive processes during a specific moment or activity (Lyle, 2003); and the present study recruited only a subset of its entire sample to participate in stimulated recall.

Moving from specifying the framework of the study’s qualitative approach, some past studies using the similar qualitative approach will be discussed to provide a critical evaluation of the research design and reference to the present study. Collins’s (2005) use of retrospective verbal report to probe learners’ internal interlanguage hypotheses about how the English tense-aspect system works and factors which may have influenced their tense-aspect interlanguage development has lent some support to the present study’s decision on using stimulated recall. She justifies the additional use of the retrospective method through the need to go from the surface investigation of learners’ tense-aspect judgment to a more in-depth level, to delve into learners’ perception of their own interlanguage hypotheses and the underlying factors. This search for a deeper answer is also why the present study chose to adopt the additional qualitative approach.

Collins’s implementation of retrospective verbal report resembles that of stimulated recall. Both use semi-structured probes to avoid directing learners to comment on their previous responses instead of revealing what actually happens in their minds at the moment during the task; interview learners as soon as the previous task ends to avoid memory decay; and use the previous task as the stimulus for probing learners’ recall of their past thoughts. These may lend references to the
present study’s implementation of stimulated recall. However, Collins’s constant prompting of learners’ recall of the same area may have made her learners realize the pattern of the interviews, and thereby expect certain responses to the semi-structured probes. Noting this potential flaw, the present study prompted learners’ recall of their thoughts on other occasions besides recasts, for example their wrong use of tense in non-recast instances and occasions that called for the researcher’s attention. Occasions that stimulated the researcher’s probing are specified in chapter 6. This way of diversifying the probing occasions may avoid focusing learners’ attention on an occasion that always stimulates further probing and from which learners can easily cast their prediction.

The use of distracters in the recall process to reduce learners’ predictability has also been used by Egi’s (2004) methodological study on comparing two verbal report techniques. She compared immediate retrospective verbal reports to stimulated recall used in probing learners’ cognitive processes underlying their language learning. The different degrees of reactivity and veridicality of the two techniques were also examined to suggest their validity issues. Reactivity is defined as the effect of a verbal report technique on learners’ subsequent learning; and veridicality concerns the accuracy of learners’ verbal reports in reflecting their cognitive processes. The results of Egi’s study showed that both immediate report and stimulated recall only had minimal effect on learners’ subsequent performances, and hence rather low reactivity. Moreover, there were no significant differences between immediate report and stimulated recall in recast distribution and recall targets. Recasts were used in the treatment sessions to reformulate the immediate report and stimulated recall groups’ production of the target as well as non-target forms as distracters. Learners’ verbal reports were correspondingly triggered at recast instances.
of the target form, recast instances at non-target forms, and correct responses. They were also similar in stimulating learners’ recall categories of target language-related episodes, other language-related episodes, comments about the task, and other episodes. However, they differed in that stimulated recall had longer length of protocol and triggered learners’ “no memory” recall; whereas immediate report triggered learners’ “no thoughts” recall.

Apart from the above findings, Egi (2004) also evaluates both the advantages and disadvantages of immediate report and stimulated recall; and through which explains the aforementioned results. First of all, immediate report is credited as allowing a brief interval between learners’ noticing and recalling points, by triggering learners’ retrieval of not-long-ago information as it is still active within their short-term memory. This is on the other hand the opposite for stimulated recall, which prompts learners’ recall after task completion. The memory decay problem is then translated to the veridicality problem of stimulated recall, because learners may reconstruct recalls due to their less active short-term memory. Furthermore, reconstruction of memory could be based on learners’ acquiring new knowledge from their double exposure to the same task stimulus as the previous treatment in the stimulated recall procedures. Immediate report does not have these memory issues because of its fresh recall and single stimulus exposure elements. Nevertheless, immediate report may incur cognitive overload when learners have to keep switching from the task to the recall focus while input processing is still engaging them.

The cognitive overload problem of immediate report may not be an issue in Philp’s (2003) implementation of immediate report; because the verbal protocol in her study only involved recall of previous recast instances verbatim, with zero qualitative probe into learners’ noticing. Despite such a departure from Egi’s (2004)
implementation of immediate report, both authors used the two-knock sounds as pre-verbal report signal. Through learners’ post-task perception, Egi found that learners did assign the specific meaning of “mistake has been made” to the knocking signal. This may reduce the validity of immediate report regardless of its memory support.

Second of all, immediate report does not have the interviewer effect because learners recall their thoughts immediately after a language episode prompted by the knocking sound, not by the interlocutor. The lack of learner-interlocutor interaction may encourage learners’ recall of their own thoughts more than their wish to well-present themselves to impress the interlocutor by covering their prior error instances. Stimulated recall however entails the learner-interlocutor interaction in its interviewing element. Notwithstanding the threat to induce learners’ polished recall in front of the interviewer in stimulated recall, the post-task non-intrusive implementation of stimulated recall may allow time for the interviewer to clarify learners’ unclear recall. This non-intrusive time allowance is lacking in immediate report. The resulted time-consuming protocol of stimulated recall is however noteworthy.

From Egi’s (2004) findings and discussion, immediate report seems better than stimulated recall in terms of no veridicality problems of memory decay and interviewer effect. Moreover, there was no reactivity problem of immediate report in affecting learners’ subsequent learning. However, the opportunities of implementing immediate report twice, each time within each of the two treatments, but just one-time stimulated recall after treatment 2 and the two posttests leading to the long time departure of stimulated recall from treatment 1, may have exacerbated stimulated recall learners’ memory decay and hence the veridicality problem. Moreover, the use of English to conduct both verbal reports, which was some learners’ L2, may have
caused the more "no memory" report of stimulated recall learners and "no thoughts" report of immediate report students. The memory decay problem of the former and cognitive overload problem of the latter may not have been the reasons. These methodological doubts of Egi's study may have obscured the research values of the two verbal report techniques.

Considering the aforementioned issues, the present study chose to use learners' L1 to conduct stimulated recall sessions and arrange the interview sessions held immediately after the immediate post-test. These may reduce the veridicality problem of stimulated recall incurred by verbalization ability and memory loss. However, the reactivity problem of stimulated recall, none in Egi's (2004) post-task administration, in affecting learners' subsequent performance may exist because the present study also had a delayed posttest for learners. The present study employed stimulated recall after the immediate posttest and before the delayed posttest to balance the reactivity, accurate report and time delay factors (Egi, 2008). About the possibly inevitable reactivity on the delayed posttest, the present study hypothesized that a three-week period of the delayed posttest away from the stimulated recall may help reduce the effect. Moreover, a statistical tool, specified in section 4.9, was employed to objectively inspect the reactivity effect of stimulated recall on learners' delayed performance.

Although Egi's (2008) later study demonstrated the non-reactivity of stimulated recall when it preceded posttests, stimulated recall is not absolutely free from projecting the reactivity effect. The nature of stimuli, type of verbalization and interaction between learners and the interviewer may have generated the non-reactivity of stimulated recall in Egi's study. Her study included the stimulated recall, stimulus, experimental control, and test control groups to investigate if recall stimuli,
verbalization or both caused the potential reactivity problem. The results showed non-significant differences among the first three groups, indicating the non-reactivity of stimulated recall, stimulus and verbalization.

Verbalization is discussed as allowing learners to rehearse their memory and thereby learn something new and change their subsequent performance, especially when the stimulus is the same or similar to the task experienced (Egi, 2008) due to reinforced exposure. Recall stimulus and verbalization are hence the two factors of reactivity. Verbalization has long been supported by Swain’s (1985, 2005) output hypothesis and illustrated in Nabei and Swain’s (2002) study as contributing to learning. Learning from stimulated recall or reactivity exhibited in the learners’ delayed posttest held afterwards was empirically shown in Egi’s (2007b) study.

Egi’s (2008) successive study advances from her two previous studies to investigate what really constitutes the reactivity of stimulated recall when it precedes both learners’ immediate and delayed posttests. The stimulated recall group participated in the interview session after engaging in the communicative activities eliciting the target structures and recast reformulation. The stimulus group only watched the video recording of the treatment process without having to recall their thoughts. The experimental control group only participated in the treatment session and the tests. The test control group attended regular classes and took the tests to scrutinize the test practice effect. The video stimulus was played to the whole class for the first two groups. Instructions were given to request learners to recall thoughts at the past event time not the recall time for the stimulated recall group. Stimulated recall learners’ thoughts were recalled and spoken to their own microphones attached, so no interviewer was involved. The stimulus group was instructed to answer some comprehension questions of the video content to their microphones.
The common non-reactivity result found in the stimulated recall and recall stimulus groups was attributed to the nature of the recall stimulus. The video stimulus recording the whole-class interaction was played to the whole class; so each learner not only recalled thoughts according to their own performance but also to their classmates'. Recall thoughts based on others' performance or needs may not contribute learning and hence reactivity. The non-reactivity of verbalization was due to the absence of an immediate interviewer audience. The learners in the study only spoke to their own microphones, not to anyone else. An audience-bearing interaction situation may on the other hand demand more processing from learners and then motivate possible learning. Egi's (2008) study inspires that the stimulated recall procedures (recall stimulus, type of verbalization and learner-interviewer interaction) may affect reactivity. The video recall stimulus played to individual learners and their thoughts recalled to the researcher's question prompts in the present study may induce reactivity on the three-week-later delayed posttest. However, the present study kept these procedures because of its intent to probe closely individuals' inner thoughts at specific moments during tasks.

Stimulated recall has mainly been shown possessing more veridicality problem than immediate report due to the memory decay and interviewer effects. The present study adopted the former because interference to the ongoing treatment task in immediate report was considered more difficult to control. The focus of the present study was on the effect of recasts, so the consistency and intensiveness of recasts as treatment were of top priority. The effects of interviewer, memory decay, acquisition of new knowledge and time consumption in implementing stimulated recall were controlled by raising non-specific question prompts, using learners' L1, following the previous task as close as possible, presenting video record of the previous task as
stimulus, and stopping learners' elaboration when necessary, to reduce harms to
validity as much as possible in the present study. These control measures will be
explained more in section 4.8.

The above studies about the validity issues of stimulated recall offer
suggestions to the present study's controlled implementation of stimulated recall.
Polio, Gass and Chapin's (2006) study additionally lends support to the present
study's use of stimulated recall to complement the deficiency of a quantitative
approach in revealing participants' perception; and through which the underlying
reasons or another perspective of the quantitative results may be unfolded. Polio et al.
examined the difference in the use of recasts between pre-service teachers and
experienced teachers in their interaction with learners in a laboratory context. They
found that the number of recasts used by the two groups did not differ, but the amount
of learners' output differed. This second finding needed to be further investigated via
stimulated recall. The qualitative findings were that experienced teachers focused
more on pushing learners to produce output and learn from their output and problems;
whereas pre-service teachers cared more about their own and learners' feelings and
the task procedures. Methodologically, Polio et al. stress the complementary value of
stimulated recall to a quantitative approach in unveiling all these differences.

The insightful feature of stimulated recall may also help suggest more
confidently if learners' surface uptakes of recasts are in fact attentive or not, when the
quantitative approach can only count the uptake instances (Polio et al., 2006). Though
stimulated recall lags behind immediate report in gaining the most active access to
learners' short-term memory when prompting their post-task recalls (Egi, 2004), Gass
and Mackey (2000) reckon that the window to learners' short-term memory in
stimulated recall is available by its recent provision of stimuli and prompting to aid

170
learners’ recall of previous events. Based on the information-processing approach, post-task stimuli and prompting may avoid cognitive burden on learners’ processing during task and relieve the memory burden involved in stimulated recall (Gass & Mackey, 2000; Nabei & Swain, 2002). Meanwhile, Gass and Mackey are cautious about the instructions guiding learners’ proper use of the stimuli and prompting to help them recall. Recall stimuli provided in stimulated recall have been claimed by Egi (2004, 2008) as potentially inspiring learners’ acquisition of new views, especially when the stimuli reflect the task process. However, Polio et al. (2006) support the use of stimuli to yield learners’ most accurate reports, which think-aloud protocol and post-task interview may lack because of zero specific stimuli or prompting. They also mention that think-aloud not only demands the same online task-switching cognitive burden as immediate report, but also language-switching between L1 and L2 from learners.

Each of the different verbal report protocols has its own advantages and disadvantages. As discussed in chapter 2, Mackey (2006) used four measures to probe learners’ noticing; they were learning journals, stimulated recall, focused L1 questions, and L2 questionnaires. The use of multiple noticing measures in one study may enhance learners’ learning or awareness of what is actually being probed even further, regardless of their counterbalancing each other’s flaws. Even when only stimulated recall is employed but in repeated cycles for the same learner, as in Nabei and Swain’s (2002) study, it may still be easy for the learner to learn from it.

Considering all the evaluations, the present study adopted stimulated recall with the control measures (see section 4.8) suggested by the relevant studies.
4.3 The Pilot Study

The following decisions made on the actual study were based on the pilot study held beforehand. The researcher first contacted the target secondary school’s principal on a personal level; and then the researcher was referred to the school English Language coordinator, who is a native-speaker from Scotland. She then asked the Form Three (S3) English Language coordinator to advertise the researcher’s attempt as a summer English workshop to S3 students. This was to return favour to the target school as well as fulfil their common practice of tutoring English to students in summer. It was held from August 1st to 25th in 2006 with 20 15-year-old S3 students from the same target school. They were having their summer vacation at that time and were in the transition to their Form Four year. The 20 students were neither from an intact class nor from random selection. All the 20 students actually volunteered to join the pilot study, except 3 of them who were asked by their Cantonese-speaking English Language teachers.

There was a mix of levels among the 20 students—8 from the advanced stream (taught by a native speaker), 6 from the medium stream (taught by Cantonese teachers), and 6 from the low stream (taught by a Cantonese teacher). The gender distribution within the 20 students was also uneven—12 females and 8 males. The imbalance of students’ English proficiency levels, teachers or classes and genders may have threatened the internal validity of the pilot study at that time. However, the pilot study tried to balance each feedback group with near-equal numbers of students from each level—2 medium-stream and 2 low-stream students in each of the three groups (corrective recast, normal recast, control), 3 advanced students in each of the corrective recast group and normal recast group, and 2 remaining advanced students in the control group.
The first half of the pilot study was the summer English workshop, teaching students the different tense and aspect forms during the first week. The lessons made use of interactive activities to teach students the daily life usage of the different forms. The study’s purpose of investigating their use of past tense was never disclosed to students at that time. The second half of the pilot study was a series of testing, feedback treatment (normal recasts, corrective recasts, content-only feedback) and stimulated recall sessions. Each of the 20 students participated in the pretest, feedback session, immediate posttest and then stimulated recall session during the second week. Lastly, the delayed posttest was held a week later. A debriefing letter, similar to the actual study’s one as displayed in appendix XI, was distributed to students together with a stationery gift as token of appreciation at the end of the pilot study.

The pilot study adopted the use of stories which mainly involved existing scenarios and characters. For example, the pre-test story was about a naughty boy asking an engineer to make a robot to replace his class teacher, so that he could be carefree about not being scolded by his class teacher about his talkativeness; the feedback session story was about a friendly boy willing to play basketball with a cat and the cat’s owner thanked him by treating him a meal; the immediate post-test story was about a helpful boy winning a prize in his Mathematics class with the help of a smart cleaning lady of the school; and the delayed post-test story was about a girl winning a cheque from her school dancing competition and gave the money to a school cleaning lady for her daughter’s medical treatment in the end. The exact cartoon-strip of each story is attached in appendix III.

As will be compared in sections 4.5, 4.7 and 4.8, the actual study had been refined from the pilot study’s experience in terms of the schedule, sample size, students’ English level, narrative story instrument and counterbalancing the order of
the different narrative parts in different sessions. All other aspects, for example the target form, target school, test, feedback and stimulated recall time allocations, instructions, and balancing students’ different feedback types across the population, in the actual study were conducted in similar manners to the pilot study.

4.4 The Study Context in Hong Kong

This section provides a comprehensive account of the background of a suburban secondary school in Hong Kong selected to conduct the present study, in order to make research replication possible. This includes 1) the public exams in Hong Kong and the school’s English proficiency level, 2) the language policy of Hong Kong and the school’s medium of instruction, 3) the Hong Kong English Language education curriculum and the school’s English Language curriculum, and 4) the school’s English Language teachers. Information given in each section below will be supported by references issued by the different Hong Kong government departments.

4.4.1 English Proficiency Level

The target school in the present study is co-educational and non-religious; and it was subsidized by the Hong Kong government at the time of the study’s data collection in March-May 2007. The overall English proficiency level of the target school can be reflected in their students’ Hong Kong public exam performance. HKCEE and HKALE are the two main public exams in Hong Kong, normally gauging secondary school students’ probabilities of graduating from secondary school and entering universities respectively (Hong Kong Examinations and Assessment Authority, 2009). In their English Language education proposal submitted in 2002 to the Quality Education Fund, which was established by the Hong Kong government in
1998 to finance school initiatives in promoting quality education (Quality Education Fund, 2009), the target school presented background information about its students’ HKCEE and HKALE results in English Language.

The HKCEE English Language passing rates of the target school’s students ranged more or less in descending order from 90% to 87% across 1998 to 2002. Moreover, less than 10% of students attained grade C or above, on the grading scale from A as the highest point to E as the bare pass adopted before 2007 (Hong Kong Examinations and Assessment Authority, 2009), in the HKCEE English Language subject from 1998 to 2002. The students’ passing rates of Use of English in HKALE from 1998 to 2001 resembled those of their HKCEE, which were also more than 80%. The target school made use of its students’ rather unsatisfactory public exams’ English Language performance to appeal to the Quality Education Fund authority that they needed to do something about their teaching curriculum or methodology to help improve their students’ English level.

Apart from using their students’ not good enough HKCEE and HKALE English Language results, the target school also quoted the rather unfavourable HKAT performance of its secondary one (S1) student entrants in 2002. The HKAT is a kind of diagnostic test for secondary schools to assess their S1 entrants’ performance in Chinese Language, English Language and Mathematics, and make necessary planning and support accordingly to enhance their S1 students’ learning in the three key areas in the coming year (Education Bureau, 2009). The target school made use of the fact that one fifth of their S1 students achieved scores under 50 out of 100 in the HKAT to draw the Quality Education Fund authority’s attention to their need of some initiatives to improve their students’ English basics.
From the target school’s background information about its students’ English Language public exam results provided in their project proposal to the Quality Education Fund in 2002, the target school’s overall English proficiency level can be attested as lower intermediate.

4.4.2 Medium of Instruction

Before the enforcement of compulsory nine-year education from primary one to secondary three (junior secondary level) in Hong Kong in 1978, English was the major medium of instruction (MOI) in secondary schools (Education Commission, 2005, p.140). Many students did not have the capability to learn in English, so schools often taught English texts to students in Chinese (Education Commission, 2005, p.140). Therefore, an International Visiting Panel suggested junior secondary 1-3 students to use Cantonese, students’ mother-tongue, to better facilitate their learning of other subjects besides English Language (Education Commission, 2005, p.140).

The Hong Kong government emphasized in one of their publications in 1997 that schools should employ mother tongue as their MOI in all academic subjects; and schools which wish to use English as their MOI must demonstrate that they “have satisfied the three prescribed criteria of student ability, teacher capability and support strategies and measures” (Education Commission, 2005, p.143). In the same year, the target school in the present study was granted the English as the medium of instruction (EMI) status, which indicated that it was qualified to use English to teach all subjects besides subjects using other languages according to the three aforementioned criteria.

However, as admitted in their 2002 project proposal to the Quality Education Fund, the target school did not receive satisfactory HKCEE, HKALE and HKAT
English Language result records from their students. This suggested that the target school’s actual English proficiency level may still need to be improved though it was given the EMI status in 1997.

4.4.3 English Language Curriculum

In the English Language Education Key Learning Area Curriculum Guide (Primary 1—Secondary 3) (The Curriculum Development Council, 2002, p.3), incorporating English teaching and learning with the aims of promoting lifelong learning and whole-person development and building school-based English Language curriculum have been recommended to schools in Hong Kong. The Curriculum Development Council (2002, p.5) identifies English Language Education as one of the Key Learning Areas for Primary 1 to Secondary 3 students. Their curriculum guide advises English Language Education to be positioned in schools’ curriculum as trying to develop learners’ English proficiency for study, work and leisure; provide them with opportunities for personal and intellectual development; extend their knowledge and experience of other cultures through the English medium; and help them overcome the challenges of the rapidly changing and keenly competitive knowledge-based society of Hong Kong (The Curriculum Development Council, 2002, p.5).

The target school’s junior secondary English Language curriculum echoes The Curriculum Development Council’s (2002) English Language Education curriculum guide in many ways. First of all, the target school adopts a school-based English Language curriculum for junior forms (Secondary 1—3), under which teachers design learning activities and teaching resources to suit different students’ interests, levels and needs. Second of all, the target school puts the Language Across the Curriculum
(LAC) (Education Bureau, 2009) methodology into practice, to let students learn different subject areas via the English medium. This appears to mirror the curriculum guide (Curriculum Development Council, 2002) that English Language Education is to facilitate students’ extensive learning or experience via the English medium.

Regarding school-based curriculum, the Secondary 3 (S3) English Language teachers of the target school for example made use of self-devised topics in addition to textbooks to teach during the study’s data collection in 2007. For instance, they used the topic, healthy living, to teach forms of adverbs of frequency, conditional sentences and adjective structure. In addition to teaching different language skills and forms through some thematic topics, the S3 English Language teachers at that time also taught different generic skills, such as critical thinking and problem solving, and different moral and civic topics. For example, they introduced integrated science and physical and aesthetic development under the theme of healthy living. Such practice seems to correspond to the curriculum guide (Curriculum Development Council, 2002) in terms of fostering students’ personal and intellectual development.

Last but not least, the target school proposed to the Quality Education Fund in 2002 that they would venture pragmatic teaching tools such as drama and extensive reading across genres and cultures in the English Language immersion program, for polishing their Secondary I entrants’ English from 2003 to 2004. Such proposal illustrated the target school’s endeavor to enhance students’ whole-person development and their interactive and communicative language skills to become usable in society in the future. All in all, the target schools’ English Language curriculum is learner-oriented, interactive and extensive beyond textbooks.
4.4.4 School Teachers

During the year when the present study was conducted from March to May in 2007, there were totally three native-speaking English teachers (NET) in the school. They were employed under the Hong Kong government's secondary school NET scheme introduced in 1997 (Education Bureau, 2009). Their major duties were providing an authentic English learning environment, coordinating and teaching English in both academic and non-academic contexts to students in the target school. The three NETs, two from Scotland and one from New Zealand, were assigned to mostly teach classes with students bearing advanced English proficiency level determined by internal assessment; and local Cantonese-speaking teachers were appointed to teach other classes with medium or low English level at that time. The three NETs in the school also assisted advanced students with preparing for international English tests to further strengthen their English proficiency as well as competitiveness.

Furthermore, NETs in the target school collaborated with local English Language teachers and other subjects’ local teachers to design school-based or learner-oriented curricula, and executed using the English medium to teach English Language as well as other content subjects. Though teachers endeavour to collaborate in constructing learner-oriented curricula and teaching, the External School Review Report administered by the Quality Assurance Division of the Education and Manpower Bureau in 2006 pointed out that teachers were recommended to give more concrete feedback on students’ work instead of just generally commenting their strengths and weaknesses. Last but not least, regular meetings and both internal and external training programmes are held for teachers to foster their teaching collaboration or exchange, plus advance their professional development.
Overall, the target school favours extensive and communicative English learning in their English Language curriculum, and obtains supports from their NETs in creating an interactive English environment for their students. However, the External School Review Report administered by the Quality Assurance Division of the Education and Manpower Bureau in 2006 suggested the target school pay greater efforts in building an English-speaking environment more quickly. This illustrated that the English environment in the target school may not be sufficient.

4.5 The Participants

This section specifies the present study learners’ characteristics in two aspects: 1) the basic identification information of the learners, and 2) the way of selecting participants in consideration of the present study’s internal and external validity. The specifications intend to provide a foundation for any possible research replication.

4.5.1 Basic Identification Information

Three intact Secondary 3 (S3) classes with 15-year-old Cantonese-speaking Hong Kong students of the target school were selected as the participants of the present study. The three intact classes were formed by screening students’ English proficiency levels based on their internal exam scores. The target school screens students into different streams or levels (advanced, medium and low) in order to allocate them to the level-appropriate English Language classes and teachers, apart from their original class divisions under school registration.

The present study first of all chose to explore S3 students’ past tense performance because S3 is the highest secondary school year within the junior forms (S1-S3). Students at the S3 level were assumed to possess the most adequate English proficiency for the study’s investigation among the junior forms. Senior forms (S4-
S7) were not considered because they are the forms which have to deal with public exams—HKCEE for S4-S5 students and HKALE for S6-S7 students (Hong Kong Examinations and Assessment Authority, 2009). Second of all, medium-stream English classes were selected because they had the largest population, 3 classes out of 5 within the whole S3 form at the time of data collection. Moreover, their English proficiency level, based on their stream assigned, was neither too advanced nor too low.

Each S3 medium-stream English Language class had 36 students and comprised students from their original classes (Class A-E). Students stayed in their original classes for other subjects, and changed to their different streams for English Language lessons. Only the advanced stream was taught by a NET, so the 3 medium-stream English Language classes were all taught by 3 different local Cantonese-speaking teachers at the time of the study.

However, not all 36 students in each of the 3 medium-stream English Language classes participated in the data collection. Prior to the actual data collection, Class 1 had 36 students, Class 2 had 25, and Class 3 had 33 agreed to participate in the study by signing a consent form in Chinese. A copy of the consent form in both Chinese and English is attached in appendix 1. During the data collection process, Class 1 lost 3 students in the delayed post-test, so only 33 students’ data were counted in Class 1; Class 2 did not lose any of the 25 students; and Class 3 had one absentee since the pretest and 1 student unable to speak in English when he/she was asked to narrate the pictures in the pretest (see appendix VI).

All together, the study’s data were obtained from 89 S3 English Language medium-stream students—33 students from Class 1, 25 students from Class 2, and 31 students from Class 3 (see appendix VI). Among the 89 students, 35 students were
assigned to do stimulated recall interviews with the researcher (see appendix VI). The way adopted to select the stimulated recall subset is described in section 4.5.2.

Although the initial choice of the three intact classes did not involve any randomization, students’ group memberships in the three different feedback groups, normal recast and corrective recast as experimental groups and content-only feedback as control group, were obtained via matching. This was to balance their different genders, teachers, class experiences and individual proficiency levels in each group. This group membership matching, implemented under both the experimental and stimulated recall methods, may help reduce threats to internal and external validity caused by the use of intact classes. Further details on internal and external validity supported by balancing group differences in both research design methods will be presented in section 4.5.2.

Regarding the students’ previous academic experience and their current courses at the time of the study, the students of the target school definitely possessed the previous knowledge of past tense. Moreover, they were having regular English Language lessons in their own stream classes during the data collection period. According to Yang, Hung and Lee’s (2000) reflections on English textbooks for Hong Kong students, the simple past is taught to students in Primary 3. Moreover, the different tenses have been introduced to all Secondary 1 students. The Hong Kong Education Bureau has initiated The Support Measures for Student Adaptation project in English medium of instruction (EMI) schools to assist Secondary 1 students with adapting to the EMI learning environment (Education Bureau, 2009). It consists of support material sets including materials helping students with basic grammar, for example the different tenses.
Furthermore, the target school’s S3 students were using *Grammarway 4* (Dooley & Evans, 1999) in their regular English Language lessons at the time of the study, which covers all the 12 different tense and aspect forms: simple present, present progressive, present perfect, present perfect progressive, simple past, past progressive, past perfect, past perfect progressive, will + infinitive, future progressive, future perfect, and future perfect progressive. This shows that the students were ready for the study’s target form—past tense.

Last but not least, concerning students’ responsiveness in English in their English as a foreign language (EFL) learning environment, most students in the target school were identified as putting efforts on presenting their own ideas in English by the External School Review Report (Quality Assurance Division Education and Manpower Bureau, 2006). However, students in the target school were observed most willing to show responses when being asked by their teachers. In other words, there was little evidence that students initiated clarifications from teachers autonomously and confidently.

### 4.5.2 Participant Selection

Which participants and how they were chosen to be examined could affect the internal and external validity of the present study, especially in the case of intact classes under the quasi-experimental design (Porte, 2002, p.40). To avoid bias within the participants themselves, the present study had attempted to keep all groups of participants on equal terms from the very beginning. First of all, only participants from the medium-stream English Language classes were considered, so as to equalize the participants’ English proficiency level as much as possible. Second of all, only S3 students were included to balance the age of the participants. Moreover, all
participants were from the same EMI school, receiving the same previous and current English Language instructions. Using students from different schools may incur the variable of different levels and exposures. Not releasing the exact focus of the present study to students in the same target school was an attempt to avoid them communicating with each other about the study’s purpose.

The present study made no attempt to investigate the relationship between learners’ different individual levels or readiness and recasts. Concerning their overall developmental readiness, the EFL students in the present study had been following regular school-based curriculum since their primary school education and they definitely had been taught the use of past tense, both regular and irregular past. They were therefore considered ready for using past tense. What they were less ready for was their proceduralized use of past tense in meaningful usage.

The present study’s learners possessed the declarative knowledge of past tense through having been instructed by their English Language teachers and textbooks. Despite this receptive instructional exposure, Hong Kong EFL learners enjoy limited productive exposure to using past tense in their spontaneous speech and hence lack the procedural knowledge of past tense. Past tense can therefore be an already-learned form (declarative knowledge) as well as a new form (procedural knowledge) for recasts to target in the present study. The present study learners’ lack of procedural knowledge of past tense originates from their EFL school learning.

The EFL education in Hong Kong normally adopts synthetic syllabuses. In Hong Kong, EFL learning is heavily based on textbook instruction in time-constrained lessons as well as exam-oriented assessment. Carless (2007) argues that this has made task-based teaching inappropriate in the secondary school context. In other words, the use of analytic syllabuses, which provide learners a communicative
context and freedom to experiment their different linguistic resources and gradually
develop the target form (Robinson, 2001a), would be a mission impossible in Hong Kong. With designated EFL syllabuses and rigidly structured textbooks prescribing teachers what and when to teach, teachers in Hong Kong often focus more on delivering preset knowledge than facilitating learners' development of communicative skills (Hu, 2005).

Undeniably, the participants bore some differences among themselves, for example their different genders, teachers, class experiences and individual proficiency levels. To balance all these, group membership matching was employed despite the use of non-randomized intact classes from the beginning. Firstly, it was fortunate that the three intact medium-stream English Language classes already consisted of near-equal numbers of male (45) and female (44) students. This had made it easy for the present study to even out the two genders across the three feedback groups (normal recast, corrective recast, content-only) and within the stimulated recall subset.

Specifically, there were 14 females and 15 males in the normal recast group and the corrective recast group, and 16 females and 15 males in the content-only control group. These had made up the group sizes of 29 in both the normal recast and corrective recast groups, and 31 in the control group. There were 20 females and 15 males within the stimulated recall subset, forming the subset group size of 35, with the not-too-different distribution of females and males. Appendix VI clearly demonstrates the number breakdown.

Secondly, to balance the participants' different teachers, different original as well as English Language class experiences and different individual proficiency levels in each feedback group, the 3 different feedback treatments were arranged to spread across each of the 3 medium-stream English Language classes. This was practically
done by matching each student from the very first one down to the last one on each
English Language class register with one feedback type in the order of corrective
recast, normal recast and then content-only. In other words, the first student on the list
was assigned to receive corrective recast, the second normal recast, the third content­
only feedback, so on and so forth. This was to try equalizing the numbers of different
feedback groups in each English Language class, in order to balance the participants’
different teacher and class experiences as well as their individual proficiency levels as
much as possible in each feedback group.

To be exact, there were: in Class 1, 12 received corrective recasts, 12 received
normal recasts, and 9 received content-only feedback; in Class 2, 8 received
corrective recasts, 7 received normal recasts, and 10 received content-only feedback;
and in Class 3, 9 received corrective recasts, 10 received normal recasts, and 12
received content-only feedback (see appendix VI). Near-equal distribution of the three
feedback types in each class has been shown.

Regarding the stimulated recall subset, near-equal distribution of the different
feedback groups was also attempted—12 students had received corrective recasts
(CR) previously, 11 had received normal recasts (NR) previously, and 12 had
received content-only feedback (CNT) previously (see appendix VI). Almost equal
distribution of students from the 3 different classes in the stimulated recall subset had
also been attained to balance the variables of teacher, class, and individual proficiency
differences—10 students (4 CR, 3 NR, 3 CNT) in Class 1 did stimulated recalls,
another 12 (4 CR, 4 NR, 4 CNT) from Class 2, and 13 (4 CR, 4 NR, 5 CNT) from
Class 3.

Although group membership matching had been sought to balance students’
different genders, teachers or classes, and individual proficiency, the variable of
having different students engaged in the feedback, test and stimulated recall sessions in different weeks and different time slots on their weekly school schedules was found not possible to control. This was due to the unavoidable constraint that the study’s data collection was conducted during the students’ regular school term, when they had their different academic and non-academic commitments concurrently.

The way to determine which student did the feedback, test and stimulated recall sessions in which time slot was done by assigning the first student on each class’s register to the first time slot of their weekly schedule. The same thing happened to other students in a sequential manner, so the second student on the list occupied the second time slot on the first day of their weekly schedule. For choosing which student to participate in the stimulated recall sessions, students assigned with less packed time slots were arranged to do stimulated recalls with the researcher in addition to their feedback and test sessions. The exact feedback, test and stimulated recall interview schedules of the different students from the 3 different English medium-stream classes are attached in appendix II.

From testing the methodology and methods in the pilot study, some limitations were found. These led to the implementation of more stringent arrangements for the actual study; for example, refined arrangements of group matching and balancing other participant variables possibly affecting the study’s internal validity as indicated above, the narrative instrument to be discussed in section 4.7, and the procedures to be delineated in section 4.8.

All the above considerations were taken to help secure both the internal and external validity of the present study. The study’s internal validity, which is whether the ultimate results reflected intrinsically the different effectiveness of the feedback types and learners’ verbalization of their perception of the feedback referred to the
treatment time, had been supported as much as possible via group matching, near-even distribution of different variables (gender, teacher, class, individual proficiency) across groups, and introduction of a control group and delayed post-test. The study's external validity or generalization to a larger population is largely dependent on the study’s internal validity (Porte, 2002, p.37); and it will be discussed in chapter 8.

4.6 The Target Form

Learners' use of past tense in narrating the different parts of a Chinese mythical story was the form being targeted by the different recast types in the present study. The use of past tense in English was selected because it is 1) a common problem among Chinese or Cantonese-speaking students, 2) a conjugated form, with the simple past in particular, which occurs early and dominantly in Hong Kong learners’ interlanguage development, and 3) a suitable candidate for recasts. The following will further narrow the scope of the present study’s target form.

The spoken mode of the narrative task in the present study made students’ performance of past tense determined not only by their knowledge and use of past tense, but also their pronunciation of the relevant verb morphological form. Therefore, learners’ phonetic output of past tense in narrating the story was a key in verifying their surface use of the form as well as the present study's data validity. However, Chinese learners’ phonetic interlanguage, which will be identified in section 4.6.1.1, may cause barriers. One way of sustaining data validity will also be specified in section 4.6.1.1. Furthermore, the reason why only the general time reference of past tense, including the simple past, past progressive and past perfect forms, was being focused instead of one of the forms differentiated by their grammatical aspectual classes will be explained in section 4.6.1.2. Extensive analyses
of English and Chinese temporal expressions with their different tense and aspect systems will also be given.

4.6.1 Chinese Learners’ Problems with Past Tense

4.6.1.1 Phonetic Problems

The use of past tense in English has been reckoned as causing difficulties to Chinese learners because of the linguistic differences between the two languages. Chang (2001) categorizes the intrinsic differences between Chinese and English in many aspects, and predicts the problems faced by Chinese learners of English accordingly.

Cantonese, the participants’ first language, is one of the Chinese dialects widely spoken in Hong Kong (Chan & Li, 2000). In terms of phonetic differences, Chang (2001) first of all recounts that Chinese learners of English have difficulties with enunciating final consonants in English, because Chinese possesses few final consonants in words. It is predicted that learners would modify the final consonant in an English word by adding an extra consonant or replace it with a glottal or unreleased stop (Chang, 2001). In the case of non-syllabic regular past where the final consonant is for example /t/ in walked /wɔːkt/, Chinese learners would probably drop the originally released stop /t/ and hence make their use of regular past indistinguishable in their speech. This was precisely the reason why the present study eliminated the count of learners’ use of regular past in their speech, to avoid ambiguous data and circumscribe the variable context for counting past tense tokens.

For the similar point on unreleased stop, Chan and Li (2000) and Chan (2010) specify it as Chinese learners’ usual problem pronouncing plosives. They distinguish spoken English and Cantonese in the sense that the former carries both released
plosives and voiced final plosives; whereas the latter does not have either of these. Released or voiceless plosives include aspirated /p, t, k/, with a burst of air released when orally articulating these consonants (Roach, 2009). Voiced plosives on the other hand cover /b, d, g/, with some vibration in the throat (Roach, 2009). Though Cantonese has all these consonants, it has neither the released nor voiced final characteristic of /p, t, k/ and /b, d, g/ respectively.

As a refinement to Chan and Li’s (2000) observational study, Chan’s (2006) experimental study rigorously investigated Cantonese speakers’ pronunciation of final single consonants. In relation to word-final plosives, she found that Cantonese speakers exhibited unreleased word-final voiceless plosives in spontaneous connected speech in the conversation task, as well as in controlled isolated speech in the word list and picture list reading tasks. These findings echoed Chan and Li’s (2000) observation of L1 interference that Cantonese final plosives are unreleased by nature. Chan (2006) additionally found the similar unreleased feature of native speakers’ final plosives, but only in spontaneous connected speech in the passage reading task and interview task and rare in isolated words in the word list and picture list reading tasks. These suggested that native speakers’ unreleased practice may have been due to their simplification strategies in spontaneous communication. Cantonese speakers on the other hand practised unreleased plosives in both spontaneous and controlled speech, due to their obligatory unreleased first language. In the case of syllabic regular past, Cantonese speakers’ pronunciation of the –ed form in wanted /wʌntɪd/, with unreleased /ɪ/, and decided /dɪdɪd/, with devoiced /d/, may not sound distinctively for listeners to detect their use of tense in speech.

Consonant clusters in the initial or final position of English words cause a similar problem to Chinese learners. According to Chang (2001), Chinese learners...
have problem pronouncing consonant clusters because Chinese does not share the same feature. They would add extra syllables or simply abandon one of the consonants out of the cluster (Chang, 2001). In the case of regular past, learners for instance would tend to drop /v/ within the final consonant cluster of /kt/ in walked /wɔːkt/; and that would again make learners’ use of regular past sound ambiguous in their speech.

Chan and Li (2000) term the adding of extra syllables as *epenthesis* and abandoning of consonants within consonant clusters as *deletion*. They not only raise that Cantonese only accommodates the consonant-vowel-consonant combination, entailing one final consonant instead of a cluster of final consonants in English, but also that alveolar consonants /l, t, d/ are often deleted in word-final consonant clusters. They therefore predict that Cantonese learners would have problem pronouncing regular past tense markers /t, d/. Under epenthesis, another problem of distribution of phonemes by Cantonese learners, an extra vowel of schwa /a/ is likely to be inserted after *walk* and before its –ed part /wɔːk/, breaking up the final consonant cluster /kt/. Either deletion or epenthesis may pose problems to native-speaking listeners’ comprehension of Cantonese learners’ use of regular past.

The phonetic ambiguity of Chinese learners’ use of regular past in speech is worth noting first because it is highly probable. The ambiguity is likely to emerge not only because of the phonetic differences between Chinese and English; but also because of the fact that consonant cluster reduction occurs legitimately in many English dialects (Bayley, 1994, 1996; Wolfram, 1985). The /t/ or /d/ final consonant within a final consonant cluster is prone to be deleted by native-speakers, due to “the phonological process of cluster reduction” (Wolfram, 1985, p.233) triggered by specific phonological environments in English. Wolfram (1985) and Bayley (1994,
1996) mention that the /t/ or /d/ final consonant of a word may be obscured by the following segment in speech if it is a consonant, and retained if it is a vowel. Wolfram (1985) adds that this phonological process inevitably happens with learners whose first language does not have final clusters. He quotes the example of Vietnamese in his study. Another example would be Cantonese in the present study.

Stemming from the above researchers' identification that learners possessing consonant cluster-free first language would experience difficulty with enunciating clusters in any case, the present study eliminated the count of past tense verbs whose surface marking in speech are heavily affected by /t/ or /d/ deletion. Bayley (1996) classifies different types of past tense verbs according to their susceptibility to consonant cluster reduction. Verbs which are not affected by consonant cluster reduction are termed strong verbs (go, was); verbs which are affected by consonant cluster reduction but retain clear surface past tense marking are semiweak verbs (left, kept); and verbs which heavily depend on final consonant cluster to give surface past tense marking are regular non-syllabics (walked, dragged).

Since native speakers' consonant cluster reduction, for example /t/ or /d/ deletion in final clusters of regular past verbs as well as monomorphemes (just, west), is a legitimate kind of phonological variation in English, it is considered as the native-like target of learners (Bayley, 1996). Bayley (1996) contends that learners who regularly engage in informal interaction with native speakers delete /t/ or /d/ from regular past final consonant clusters more often than those classroom learners. He elaborates that such target omission of final /t/ or /d/ does not signify learners' grammatical unmarking of the form; rather, it illustrates learners' approaching the target use. However, there still exists the possibility that learners simply do not mark any tense at all by retaining the base form (Bayley, 1996).
The possibility of unmarking the underlying form in speech would more likely emerge as a case for Chinese learners, because of the tenseless trait of Chinese in expressing temporal semantics. Moreover, unlike Bayley’s (1996) Chinese learners living in the U.S., the Hong Kong Chinese learners in the present study did not have the exposure to a native English-speaking environment. They may therefore become less able to internalize the target-like /t/ or /d/ deletion in regular non-syllabics and realize actual marking of the underlying form at the same time.

Although the above early studies have shown that /t, d/ deletion in final consonant clusters is universal across English dialects, recent sociolinguistic studies put emphasis on the different constraints favouring and disfavouring /t, d/ deletion. Tagliamonte and Temple (2005) and Smith, Durham and Fortune (2009) review the different phonological and morphological effects on native speakers’ coronal stop deletion or /t, d/ deletion. The phonological constraints include following and preceding phonological segments; and the morphological constraint is the grammatical category of the relevant morpheme. Following phonological segments which favour /t, d/ deletion in the order of strength are obstruents, e.g. I have a present for you; liquids, e.g. I am going to find Lily; glides, e.g. We are experiencing wild wind; and pause, e.g. I haven’t finished. Preceding phonological segments which favour /t, d/ deletion in the order of strength are stops, e.g. He dropped me off at the bus stop; nasals, e.g. I am meeting my friend Judy; and other fricatives, e.g. We left Mary there. Morphological class predicts that monomorphemes, e.g. just, favour /t, d/ deletion more than semiweak verbs, e.g. left (irregular past with vowel change and /t/ as final consonant); and semiweak verbs favour /t, d/ deletion more than regular past tense, e.g. walked. This ranking of semiweak verbs and regular past tense is the exact opposite of that of Bayley (1996). This is mainly because sociolinguists do not
specifically take the ease of detecting past tense tokens in speech into consideration. These different factors suggest that /t, d/ deletion in final consonant clusters may not be categorical, but variable depending on the different aforementioned phonological and morphological contexts.

Although some of the above contexts disfavour /t, d/ deletion, the present study circumscribed the variable context of counting past tense tokens by not counting regular past. This is because the favouring and disfavouring constraints of /t, d/ deletion identified in sociolinguistics apply predominantly to English native speakers, not second or foreign language learners of English. Moreover, the complexity of the different phonological constraints requires listeners' sharp discrimination of speakers' /t, d/ deletion or not in real time. To circumvent erroneous findings, the present study resolutely excluded regular past.

As indicated that Chinese learners would have problem pronouncing regular past distinctively in their speech, irregular past form, which stands as individual lexical items (Ellis, 1987) and is phonetically salient without easily being affected by learners' interlanguage as well as the phonological process in English (Bardovi-Harlig, 1999; Bayley, 1994, 1996; Salaberry, 2000; Wolfram, 1985), has become the key target of the present study. Irregular past is referred as past tense hereafter. Learners' irregular past tokens and obligatory contexts for irregular past forms were considered in the ratios used to measure learners' use of past tense. Details on ratios will be covered in section 4.9. Ellis et al. (2001) also found that focus on form targeting lexical form triggered higher uptake than grammar form. However, Cantonese learners' L1 interference may still affect listeners' detection of their accurate use of irregular past.
Chan and Li (2000) identify the English phoneme inventories which Cantonese lacks. One of the examples related to past tense is long versus short vowels. Phonemically, Cantonese carries mostly short vowels (Chan & Li, 2000; Chan, 2010), and does not carry any long vowels except /a:/ as in /sa:n(1)/ meaning hill in English (Chan & Li, 2000, p.9). Accordingly, Cantonese learners may not be able to distinctively pronounce the long vowel /ɔː:/ in irregular past caught /kɔː:t/ for example; they may most likely use its short vowel counterpart /o/ for strategic substitution (Chan, 2010). Phonetically, as Chan (2010) found in her advanced Cantonese speakers’ data, neutralization of the English long and short vowels may result. Chan (2010) reckons that this is due to the similarities between some English long vowels and the long allophones of some Cantonese short vowels, rather than solely the absolute differences between English long vowels and Cantonese obligatory short vowels. However, it is noteworthy that Chan’s (2010) study involved advanced learners who may not have exhibited the common or systematic interlanguage phonology of Cantonese speakers in general.

No matter whether it is the impact of the absolute differences between the two languages or the similarities, Cantonese speakers’ distinguishable pronunciation of some irregular past may be affected. Nevertheless, this could be remedied by relying on the constrained context of learners’ elicited speech data from a predetermined narrative story (Stibbard, 2004). Stibbard (2004) encourages the use of spontaneous speech data rather than isolated ones to fairly investigate learners’ pronunciation. However, this may run the risk of having learners’ variation of pronunciation in spontaneous connected speech. The constraining of lexical choices in elicited data may solve the problem of learners’ confusing output. Meanwhile, Stibbard cautions that learners’ guessing may be resulted; and too much reliance on top-down
processing without any feedback intervention may weaken the validity of data. The combination of elicited speech and consistent feedback intervention in the present study may be a resolution to such a worry.

Irregular past was also the only target form in the pilot study. Throughout the different test and feedback sessions, learners exhibited the use of both past and non-past forms. There were occasionally some regular past-sounding forms, with learners’ adding an extra syllable resembling /dad/ for the –ed part of walked for instance. This, adding /dad/ for the –ed part of walk /wɔːd/, and the above prediction made by epenthesis, adding schwa /ə/ after walk and before the –ed part of walk /wɔːk/, may illustrate Stibbard’s (2004) point on learners’ variation of pronunciation of a single word in spontaneous connected speech. It would however be difficult to take these unstable regular past-sounding forms as learners’ instances of using regular past, though it is likely that these kinds of forms are Chinese learners’ interlanguage pronunciation of regular past. Therefore, same as the pilot study, the actual study eliminated the data ambiguity involved by counting only irregular past forms.

From learners’ instances of using regular past-sounding forms in the pilot study, the actual study learned that researching on discrete linguistic items may not necessarily have deprived learners of the opportunity of learning the global use of past tense for all verbs. Learners’ seeming attempt at regular past forms may have demonstrated their extensive use of past tense, regardless of the study’s exclusive focus on irregular past. Moreover, in section 3.3 of chapter 3, researchers’ position that there is an interface between explicit and implicit knowledge suggests learners’ development of extensively using the target form across different instances after intensive exposure to particular formal input or feedback. This may help clear the
worry that input or feedback on discrete linguistic items restricts learners’ use to the individual items.

4.6.1.2 Grammatical Problems

The use of past tense in English is also a challenge to Chinese learners grammatically, because Chinese is a language without verb conjugation (Chang, 2001; Huang, Yang & Tickoo, 1999; Yang & Huang, 2004; Yang, Huang & Lee, 2000). Chinese expresses time reference by means of temporal adverbials, chronological contexts, or indirectly through aspect markers, instead of inflecting verbs. Tense morphology in expressing time in English is generally acquired later by Chinese learners (Bardovi-Harlig, 1999; Cai, 2003; Chang, 2001; Yang & Huang, 2004), and considered difficult by Chinese learners because of their use of time adverbials in L1 overshadowing tense use in L2 (N. Ellis & Sagarrà, 2010a, b). This is probably why Chinese learners often produce errors when changing verbs into either their regular or irregular past form in English (Chang, 2001).

Bardovi-Harlig (1999, 2000) notes that learners seem to undergo developmental stages of first relying on chronological order, next adverbials, and then verb morphology in expressing temporality. Although learners in general are predicted to eventually employ verb morphology in conveying time meaning, Yang and Huang’s (2004) study discovered that their Hong Kong learners, coming from primary school to university levels, still attached the use of tense with temporal adverbials. This finding suggested that Chinese learners’ first language, which depends on temporal adverbials instead of tense as one of the means to refer to time, largely influences their use of tense in English.
Second or foreign language learners’ acquisition of tense in English has been predicted by the *aspect hypothesis* (Andersen, 1991) and *discourse hypothesis* (Bardovi-Harlig, 1992a, b, 1994, 1995) to follow a certain path. Under the aspect hypothesis, learners exhibit early use of verbal morphological encoding only according to the inherent aspect of lexical items, not tense or grammatical aspect. Specifically, learners will first mark past tense in verbs carrying the lexical aspect of achievement, then accomplishment, and then activity, and lastly state. These are the four categories of lexical aspects classified by Vendler (1967). According to Vendler (1967), *achievement* verbs, e.g. *find*, are punctual, under which actions are gone once they happen; *accomplishment* verbs, e.g. *build*, are durative with a clear starting and ending point; *activity* verbs, e.g. *run*, are durative, dynamic but without a clear endpoint; and *state* verbs, e.g. *like*, are non-dynamic and unchangeable. Achievements are the early ones being marked in learners’ acquisition process, because their meanings of complete and instantaneous actions match the perfective grammatical aspect of past tense, the simple past more precisely.

Howard’s (2004) study finding substantiated the aspect hypothesis by showing that achievement verbs favoured past time marking. Moreover, the finding also supported the discourse hypothesis by showing that learners’ past tense marking occurred more often in the foreground, which involves the use of dynamic verbs in moving events forward along the narrative timeline. It was less in the background, which involves the use of state verbs in providing related information to support the foreground. His study therefore suggests that both the aspect and discourse hypotheses interact to project multiple effects on learners’ past tense interlanguage. The discourse hypothesis will be further elaborated in section 4.7.
From the aforementioned terms of lexical aspect and grammatical aspect, their different meanings can clearly be noted. As explained in the aspect hypothesis, lexical aspect is the situational or inherent aspect of a lexical item that expresses an action (Comrie, 1976). In other words, it is about the type of event under focus. Grammatical aspect is on the other hand the viewpoint aspect of an event encoded through inflectional morphology (Smith, 1997). For instance, if an event is viewed in its entirety, then perfective marking, most notably through using the simple past, will result. If an event is however viewed as ongoing, then imperfective marking, most prominently through using the progressive –ing form, will result. Therefore, according to the aspect hypothesis, the lexical aspect of an event is being prioritized by learners over the tense or grammatical aspect when engaging in morphological inflection. In such a case, tense and aspect can be considered as separate in English. However, this may not be the case for Chinese.

Chinese makes use of lexical markers to denote both completed and past events. According to Cai (2003), le, zhe and guo are the main grammatical aspect markers in Mandarin Chinese. Yang, Huang and Lee (2000) also provide the Cantonese Chinese grammatical aspect markers equivalent to le, zhe and guo, which are jo, jyu and gwo respectively. In this sense, Chinese shares the similar, if not exactly the same, use of grammatical aspect with English. Cai (2003) particularly draws a connection between the aspect marker le in Chinese and the perfect aspect and simple past in English. She argues that the “verbal le does not locate events in past, but its major function of providing an entirety view enables it to be used in talking about past events, like the counterpart of English simple past tense (perfective past)” (p.53). Meanwhile, Cai (2003) admits and Yang, Huang and Lee (2000)
indicate earlier that there are significant differences between English past tense marking and Chinese / Cantonese aspect marker le / jo.

In general, the use of le / jo is more restrictive than English past tense marking (Cai, 2003; Huang, Yang & Tickoo, 1999; Yang, Huang & Lee, 2000). Unlike English, there is no deictic context for the aspect system in Chinese. Chinese employs temporal lexical expressions as the deictic center, such as yesterday or last month, but the perfective aspect markers le / jo mainly serve to indicate completeness of events. No reference of events to any deictic context is made. However, as shown in the following example provided by Huang, Yang & Tickoo (1999), le / jo is sometimes used to relate one event to another.

\[
\begin{align*}
wo\ chi-le & (perfective) fan & jiu huijia (Mandarin) \\
I & eat-jo (perfective) meal & then go home (Cantonese in English translation) \\
I & went home after I had finished my meal. & (English in past tense) \\
Or & I will go home after I finish my meal. & (English in the future time)
\end{align*}
\]

Le / jo describes the completed state of eating the meal as the first event; it at the same time relates this first event to the subsequent event of going home. However, it does not entail any specific temporal location; no deictic context is involved in other words. Thereby, the two Chinese sentences can be used concurrently to refer to the past and the future; but the two time references have to be expressed in two separate sentences in English. Confining to the main meaning of completeness without the accompanying deictic context is how restrictive le / jo is in Chinese. Meanwhile, the past and completed meanings of le / jo in Chinese share similarity with the conjoined tense and grammatical aspect systems in English.

The present study chose to consider the grammatical aspect interpretation of the English aspect system, instead of the lexical aspect one. This is because of the comprehensibility of the former in the Chinese system, and evidence of the effect of
the former on Chinese learners’ use of past tense in English. Bayley and Langman
in English to investigate if there is any group and individual convergence of results.
They found that Chinese learners’ past tense marking was significantly affected by
the grammatical aspect of English and the perceptual salience factor on both group
and individual bases across different proficiency levels. Specifically, the Chinese
learners marked perfective verbs or actions more likely than imperfective ones, and
marked perceptually salient verbs more than non-salient ones.

Both Wolfram (1989) and Bayley (1994) establish saliency hierarchy models
for ranking which past tense forms are more likely to be marked than others.
Wolfram’s model puts irregular past on the higher rank than regular past because the
former appears more distant from the root form. Wolfram’s saliency ranking of the
different irregular past forms follows as suppletives (e.g. go / went or am / is / was)
being the most salient and easily marked, vowel change plus suffix (e.g. leave / left),
internal vowel change (e.g. come / came), modal verbs, replacives with no change of
vowel but affixation of /t/ or /d/ (e.g. have / had or send / sent), and regular past being
the least salient and easily marked.

Bayley’s (1994) model provides a different and more detailed ranking of past
tense marking. He predicts that “the more salient the difference between the present
and past tense forms, the more likely a past-reference verb is to be marked” (p.161).
Accordingly, suppletives (e.g. go / went or am / was) would most likely be marked by
students among other forms of past tense marking; and it can be inferred that
students’ use of suppletives would appear most apparent in their speech and clearly be
detected by the researcher in the present study. Other forms in descending order of
saliency along the hierarchy are verbs which involve a vowel change and the
affixation of /t/ or /d/ in their past form (e.g. leave / left); verbs which involve vowel change and final segment deletion with /t/ or /d/ affixation as replacement (e.g. bring / brought); verbs which involve vowel change in their past form (e.g. come / came); copula verbs which involve one common segment with their present form (e.g. is / was, are / were); replacives which involve no change of vowel but affixation of /t/ or /d/ (e.g. have / had or send / sent); regular nonsyllabics (e.g. play / played); regular syllabics (e.g. want / wanted); and lastly modal verbs.

Both saliency models basically coincide on ranking suppletives and vowel change plus suffix /t/d above internal vowel change. However, Bayley adds the entry of vowel change, deletion in final segment and suffix /t/d to replace the deleted final segment, and places it above internal vowel change. Apart from this addition, Bayley also approaches a more phonetic-borne analysis and splits the past form *was* into two categories—suppletive, the past form shares nothing with the present form, e.g. *am* versus *was*; and copula verb, the past form shares one common segment with the present form, e.g. the final consonant of *is* and *was*. The second category is placed after the internal vowel change entry. Another phonetic-borne analysis is adopted for ranking the regular past. The regular past is divided into two categories based on their syllabic features in speech. Nonsyllabics are considered more salient than syllabics. This ranking may be doubtful because the marking of nonsyllabics in speech would become more obscure than syllabics due to the universal phenomenon of /t/d deletion in final consonant clusters. They are both placed at a lower rank than replacives.

The last difference is Bayley’s placing modal verbs at the bottom of the hierarchy, whereas Wolfram considers modal verbs more salient than replacives and regular past. The superiority of modal verbs over the other two entries in Wolfram’s model or the inclusion of modal verbs in both Wolfram and Bayley’s past marking
saliency models is inconceivable. This is because the tense forms of modal verbs are not definite; in other words, certain forms are the tense forms of modal verbs depending on usage and meaning. As noted by Swan (2005), modal verbs do not normally have past forms, except when would, could, should and might are used as the past tense forms of will, can, shall and may respectively under the occasion of reported speech about past events. Moreover, would, could, should and might by themselves can be in the present tense conveying different degrees of certainty, possibility, or obligation. Past meanings of modal verbs entailing these different degrees of tentativeness are usually expressed by a modal verb followed by a perfect infinitive (have + past participle). Due to these different usages and meaning occasions, the tense forms of modal verbs can be represented in different ways. The indefiniteness involved may make it difficult to determine the saliency level of the past form of modal verbs as well as the specific tense form of modal verbs.

Based on the above two saliency models and the queries derived, the present study decided to eliminate replacives, regular past, and modal verbs in its coding system and hence dataset. Verbs which have both their base form and past form remained the same were abandoned because of their tense ambiguity (e.g. put, read). All the rest of the irregular past forms, suppletives, vowel change with suffix t/d, vowel change with deletion of the original final segment but addition of suffix t/d, internal vowel change, and copular verbs, were considered as past tense tokens used by learners in the present study. Besides these different conjugation types of irregular past, tokens with irregular past preverbal markers be (e.g. was take, was taken, was taking, was took) and auxiliary structures do (e.g. did take) were also counted, because they may instantiate Chinese learners' interlanguage.
Be is prevalently used by Chinese students because of their instructional context. Cai (2003) explicates in her result analysis that “be is the earlier English verb [the Chinese] students [in her study] started to learn, [and] were [therefore] very familiar with it. In [her] students’ learning of English simple past, the textbooks and their exercises all emphasized the function of be in forming simple past forms” (p.97). Thereby, learners’ prioritized use of past tense marking in the form of verb-raising over inflection or conjugation may be resulted. In Ionin and Wexler’s (2002) study with Russian children learning English, learners’ tense marking is realized in the form of verb-raising through adding be auxiliary in front of a main verb, e.g. He is help people, rather than surface inflection on the main verb to produce helps. They conclude that their learners are only deficient in surface morphological marking; they are not impaired at the tense grammar feature. Learners’ interlanguage surface marking may be worth considering as the presence of their underlying tense usage.

Both the lexical and grammatical aspect systems in English predict that learners mark verbs or actions with perfective or completed meaning more likely than imperfective meaning in their early stage of past tense marking. They only indicate learners’ certain developmental stages instead of their holistic use of tense across verb types, which is learners’ ultimate goal of acquisition. The aspect hypothesis for instance may be defective in explaining cases where learners concurrently show tense marking in state verb copula be, as in Haznedar’s (2007) study with a Turkish child learner of English. Section 4.9 will further justify the present study’s choice of not using learners’ interlanguage developmental stages and its choice of using learners’ emergent interlanguage form marking as indices to measure their use of past tense.

Instead of delving into learners’ grammatical aspect marking in English, the present study focused on Chinese learners’ generic past tense marking in the simple
past, past progressive and past perfect. This is primarily because being tenseless is one prominent difference of Chinese from English (Bayley & Langman, 2004; Chang 2001; Yang, Huang and Lee, 2000). Therefore, learners' past tense marking in the past progressive and past perfect was determined by the marking of auxiliary verbs be and have respectively; the progressive form and past participle attached were not considered. Past marking in have was however not counted because of its being a replacive.

All in all, due to Chinese learners' major difficulty with tense marking, the present study selected learners' problematic attempts at past tense as the target of recasts. Learners' past tense marking tokens, not necessarily the accurate or error-free usage in their spoken narratives, were counted as their attempts at past tense. For example, was take, was taken or was taking were all counted as learners' irregular past tense attempts, even though they were inappropriately used in contexts where the simple past was actually the accurate form. Coding rules of past tense tokens and what counts as past tense for the purposes of data analysis are detailed in appendix VIII. In other words, the different types of recasts targeted zero surface sign of learners' attempts at irregular past marking in their treatment sessions.

Apart from the difference in time referencing, adjectives and verbs are often used identically and yet grammatically in Chinese (Chang, 2001). For example, I busy with work is equivalent to I am busy with work in Chinese both grammatically and semantically. In the case of an English sentence with adjective predicate, Chinese learners would make it verbless when their first language casts a transfer effect (Chang, 2001). Because of this valid use of grammar in Chinese, verbs and hence verb tenses would sometimes be absent in their speech.
4.6.2 Hong Kong Learners' Familiarity with the Simple Past

Although past tense marking is one distinct feature of temporal semantics in English compared to Chinese and it is hypothesized to be unfamiliar to and hence difficult for Chinese learners, it has been found occurring in an early stage as well as in significant amount in Hong Kong learners’ English textbooks. The early and significant occurrence of past tense in students’ textbooks may serve to jumpstart their interlanguage development. In their review of the Hong Kong primary school students’ tense and aspect production and their English textbooks, Yang, Huang and Lee (2000) first mention the internal factors of L1 influence, tense-aspect form complexity and lexical aspect role in affecting learners’ tense and aspect acquisition. They then focus on the external factor of English textbooks, which are nearly students’ sole input of English tense and aspect under the Hong Kong English Language education system.

Yang, Huang and Lee (2000) found that the simple past was a verb form appeared as early as primary three in the English textbook series their Hong Kong learners used; the two conjugation forms being introduced in the primary one and two textbooks were the simple present and present progressive respectively. Among these three verb forms, the simple present (48%) constituted the most distribution percentage across the English textbook series from primary one to six; and simple past (44.5%), with more or less the same percentage, came next in terms of its distribution percentage across the textbook series. The past progressive and past perfect were the two conjugation forms emerged later in primary five and six textbooks respectively; and each bore less than 1% of distribution rate across the textbook series.
The distribution figures could have been the reason why Yang, Huang and Lee (2000) obtained the dominating figures of students’ use of the simple present and simple past, 47% and 51% respectively, in their written narratives. Less than 2% of each form, the present progressive, past continuous and past perfect, was however used by students in their narratives. Yang, Huang and Lee’s (2000) review of the primary school English textbook series and finding of their primary school students’ production showed that the simple past was the verb form which appeared early and dominant in Hong Kong students’ education. Hence, the simple past was predicted to be the most apparent token signifying learners’ generic use of past tense in narratives in the present study.

Especially when narrative was used for the elicitation of tense and aspect forms, Yang, Huang and Lee’s (2000) Hong Kong learners demonstrated their dominant production of the simple past. The simple past is inferred to be most preferably used by learners for achievement verbs (Vendler, 1967) which share the same inherent feature as the simple past of being punctual, instead of durative, to express each complete event in a sequential narrative before the other. Because the present study employed story narrative to elicit learners’ production, the simple past was predicted as the apparent token of past tense. This prediction is based on both the aspect hypothesis (Andersen, 1991; Andersen & Shirai, 1994, 1996; Li & Shirai, 2000; Bardovi-Harlig & Bergstrom, 1996; Bardovi-Harlig, 1998, 1999) and the discourse hypothesis (Bardovi-Harlig, 1992a, b, 1995, 1998, 1999) proposed for determining learners’ use of tense and aspect in their interlanguage development.
4.6.3 Candidate for Recasts

As discussed previously, the irregular past tense has been chosen for recasts to target in the present study. The rationale behind is that irregular past is both physically salient (Goldschneider & DeKeyser, 2001) and meaning-bearing. It is physically salient in terms of sounding both phonetically and phonologically distinguishable (Bardovi-Harlig, 1999; Bayley, 1994, 1996; Ellis, 1987; Salaberry, 2000; Wolfram, 1985). It is meaning-bearing because it conveys temporal sense and therefore contributes communicative value. Long’s (2007) argument may confirm the appropriate choice of irregular past for recasts in the present study. He specifies that recasts would be more effective in channeling learners’ attention to the linguistic form under focus when the form is both physically salient and meaning-bearing.

Williams and Evans’s (1998) specification may overall justify the grammar of time expression as a suitable candidate for recasts. They state that “…those that differ in non-obvious ways from the learners’ first language… [are]… likely candidates for effective focus on form…” (p.140). Though irregular past with salient verb conjugation in English is considered as formally distinct from the Chinese way of expressing time via adverbials for example, the alternative of using time adverbials in English can be interpreted as non-obviously different from the Chinese means of time expression. English uses past tense through verb inflection with or without adverbials. Chinese learners however usually express time meaning in English with the mere use of adverbials, without applying any verb inflection. Under this interpretation, the means of English time expression can appear non-obviously different from the Chinese way.

Taking learners into consideration, a linguistic form which is developmentally ready for learners would be a suitable candidate for recasts (Han, 2002). In Han’s
(2002) study on the effectiveness of recasts in facilitating learners’ tense consistency in their narratives, learners’ developmental readiness is argued as one of the factors determining learners’ attention to the form being targeted by recasts. As delineated in section 4.5.1, learners in the present study had previous knowledge of past tense. Moreover, past tense has been identified in Yang, Huang and Lee’s (2000) primary school English textbook review as already introduced in Hong Kong students’ primary education. Learners in the present study were therefore considered developmentally ready for irregular past and predicted to be able to attend to it in recasts. As discussed in chapter 3 and section 4.5.2, the procedural knowledge of past tense usage in speech can however be considered as novel to Hong Kong EFL learners.

The present study therefore investigated recasts, which are a type of interlanguage-sensitive feedback and can be contrived as proactive intervention through its consistently targeting a particular form arisen from learners’ elicited speech. Jointly with previous discussions about recasts, the present study is motivated by: 1) the developmental-insensitive teaching approach in Hong Kong, and 2) the corrective function of recasts to non-disruptively and timely address one particular target form to learners, through framing focus on form within learners’ same meaning context and immediate and consistent contrast made to their individual interlanguage forms (Doughty, 2001, 2003; Long, 1996, 2007).

To experiment the use of recasts in realising a more development-sensitive approach to respect EFL learners’ internal syllabus, the present study utilized task as the instrument to elicit learners’ interlanguage forms for recasts to correct and closely react to their current development. The present study adopted Robinson’s (2005) rationale and designed its task complexity as requiring learners’ use of past tense, to
encourage their seeking help from recasts. Regarding form accuracy and complexity as discussed in chapter 3, specification of choosing which to be the present study’s target measure is needed, so that corresponding task control can be arranged to favour the desirable outcome. The present study’s EFL learners were of lower intermediate with very little opportunity of English exposure. Therefore, it would be more realistic to consider learners’ emergent attempt of past tense (grammatical complexity) as the dependent variable, rather than their accurate use of it. Skehan (1998) defines complexity as learners’ risk-taking attempts beyond absolute error-free accuracy in speech production. Moreover, Yuan and Ellis’s (2003) study demonstrated that pre-task planning benefited grammatical complexity and online planning fostered accuracy in addition. Although Romanova (2010) found that online planning outperformed pre-task planning in lessening learners’ online cognitive burden and thereby facilitating their learning from recasts, the present study adopted pre-task planning. The benefit of pre-task planning seems more focused in facilitating complexity; and online planning may incur disruption to task completion potentially caused by speech monitoring.

4.7 Instrument

This section introduces the instruments used in the feedback and test sessions as well as in the stimulated recall session of the present study. The present study responded to the need of using more accurate measurement of learners’ degree of automaticity by adopting the use of spoken narratives. The spoken narrative story, as delineated in section 4.7.3, was structured by a fixed storyline and corresponding pictures in tight sequence. This was to attempt to resolve Ammar’s (2008) worry that free production tasks may not impose enough time pressure on learners for
investigating their automaticized use of form. Moreover, the present study employed stimulated recalls to probe qualitatively, instead of quantitatively like the measurement of reaction time in Ammar’s (2008) study, into learners’ automaticity of using the target form.

Spoken narratives on the different parts of a Chinese mythical story were the instruments adopted in the feedback and test sessions. The following will first of all make use of the discourse hypothesis to explain how the discourse genre of narrative elicits the use of past tense. Then, the precise way of implementing the narrative instrument in the study will be specified, to show how reliability and validity were sustained at the same time. Moreover, the study’s choice of carrying out the narrative instrument in the spoken mode rather than the written mode will be justified. As discussed in sections 4.2.1 and 4.2.2, some studies implemented recasts during learners’ narrative output and others through learners’ interaction with their interlocutors. The present study’s use of the former instead of the latter will be explained. Lastly, the use of stimulated recall procedures will be mentioned to emphasize that they were crucial in eliciting learners’ responses about their cognitive thoughts.

4.7.1 The Narrative Discourse

The narrative discourse was selected to be the study’s instrument. The choice was based on the likelihood that learners’ use of verb tense and aspect in English can be elicited, when expressing the time sequence of consecutive event happenings characterizing a narrative (Bardovi-Harlig, 1992a, b, 1995, 1998; Dahl, 1984; Yang & Huang, 2004). In particular, the simple past has been found the dominant form elicited to narrate the foreground of a narrative; foreground often leads event
happenings to move forward along a sequential timeline (Bardovi-Harlig, 1992a, b, 1995, 1998, 1999; Dry, 1983; Hopper, 1979). Other past forms as well as non-past forms are often used to construct the background of a narrative; background often provides supplementary information about foreground (Bardovi-Harlig, 1992a, b, 1995, 1998, 1999; Hopper, 1979).

The relation between tense and narrative structure can be drawn by the aspectual meaning of the grammatical form employed and the function of the particular narrative structure. The complete and non-durative nature of each event following each other within a narrative exactly matches with the completeness and punctuality meaning of the perfective past, or specifically the simple past (Bardovi-Harlig, 1992a, b, 1995; Dowty, 1986; Hopper, 1979; Reinhart, 1984). Regarding background, the simple past, other past forms (the past progressive, past perfect) and non-past forms (the simple present, present progressive, present perfect, future) are all possible forms used to express information prior to the main events driving the story’s timeline, outcome of the main events in the end, or evaluation of the main events at any point (Bardovi-Harlig, 1992a, b, 1995, 1998, 1999; Hopper, 1979).

Although particular tense and aspect forms have been identified to be probably used by learners according to the foreground and background division of a narrative (Hopper, 1979), native speakers’ use of past tense in both foreground and background is learners’ target acquisition as their interlanguage development advances (Bardovi-Harlig, 1992a, b, 1995, 1999). During learners’ interlanguage development, they would firstly acquire the use of the simple past in foreground and then later generalize it to background (Bardovi-Harlig, 1992a, b, 1995, 1998). Predictably, the use of a foreground-background-structured narrative can elicit the present study low-
intermediate learners’ most attempt of past tense in expressing the foreground; and thereby provide a past tense context for recasts to target.

4.7.2 The Choice of Elicited Narrative

The decision made on the choice of the present study’s instrument was the use of elicited narrative instead of a controlled grammar exercise, personal narrative or conversation interaction. As discussed in section 3.2 of chapter 3, a controlled grammar exercise may constrain learners’ spontaneous use of the target form and the opportunity to probe their real-time application of the form, especially when the form is already learned but short on learners’ usage. Elicited narrative based on tightly pictorially-structured cartoon strips was preferred, because learners’ cognitive load of organizing loosely structured elements may otherwise be increased and their formal focus may be severely hindered (Foster & Skehan, 1996; Skehan & Foster, 1997, 1999; Tavakoli & Foster, 2008, 2011) as a result. In addition to giving a tighter structure, the use of elicited narrative is also more ideal than personal narrative, which is looser in structure, in terms of involving easily identified action sequences, namely foreground, to elicit the likely use of the simple past. As discussed in chapters 2 and 3, Lesser (2008), Ellis (2009a) and Philp and Tognini (2009) contend that a production task has to be meaning-based and contrived in a way to demand learners’ grammatical encoding of the relevant meaning. Narrative may exactly help balance the test of what learners know and what they can do with their L2 (Ellis, 2005a).

Elicited narrative is also favoured over conversation interaction for recasts to perform non-interfering correction. This is because elicited narrative can control the potential ambiguity in learners’ meaning and avoid interlocutors from modifying learners’ original meaning (Hauser, 2005). Hauser (2005) argues that the main
reformulating function of recasts may be distorted in the setting of conversation interaction. In the case of a learner’s incomplete and therefore ambiguous propositional meaning in his/her error-bearing turn, the interlocutor may have to determine the learner’s propositional meaning based on his/her own interpretation before reformulating the error form within the context. Such reconstruction of learners’ meaning may then distort or contradict the non-interfering function of recasts in reformulating learners’ errors without changing their original or intended meaning during ongoing interaction (Long, 1996, 2007; Long & Robinson, 1998). Thereby, learners may become confused with the corrective purpose of their interlocutors’ recasts, whether it is a meaning or formal modification. Chapter 3 has identified another complexity involved in dialogic interaction, which is learners’ elliptical responses to interlocutors’ clarification requests occurred in interaction. All in all, the present study used monologic and elicited narrative as the instrument.

4.7.3 The Present Study’s Narrative

The narrative story implemented in the present study was based on a traditional Chinese mythical story, well-known as “Archer God’s Shooting Ten Suns and Chang Or’s Flying to Moon”. It was adapted for the purpose of lengthening it into four different scenes and including different foreground and background elements in each scene. It was assumed that the universal foreground-background structure of the narrative discourse (Hopper, 1979) can thereby be formulated for learners’ target acquisition of using past tense for both elements (Bardovi-Harlig, 1992a, b, 1995, 1999). Furthermore, providing students with the exposure to the same story context all along can promote learners’ familiarity with the task. Familiarity with task
material has been organized by Skehan (1996) as one of the cognitive factors constituting task difficulty. To ease potential task difficulty, "Archer God’s Shooting Ten Suns and Chang Or’s Flying to Moon", which is a well-known traditional Chinese mythical story behind the annual Chinese mid-autumn festival, was chosen as the narrative story. Moreover, splitting it into four different scenes for learners to narrate one of them in their pre-tests, feedback sessions, immediate post-tests and delayed post-tests respectively was to lessen learners’ cognitive burden potentially caused by orienting to a brand new story context each time. Thereby, learners’ cognitive space could be freed for formal focus in both their narratives and recast treatments, to facilitate research reliability and data validity.

The idea of selecting a traditional Chinese mythical story was derived from the pilot study. Each cartoon-strip story used in the pilot study also had different foreground and background elements, typical of a narrative structure. The results however showed that nearly all learners in the pilot study used base form or present tense for narrating each story; possibly because the scenarios and characters were all near-authentic ones and using historical present to narrate stories is legitimate among native speakers when immediacy is pursued. Hence, the present study changed to the use of a mythical story with fictional characters and scenarios to steer learners towards using past tense for historical contexts, especially when learners may not have the native-like capacity to employ the generic and automatic use of past tense for narratives. Moreover, the use of a brand new story in each test and feedback sessions in the pilot study did not support the task familiarity factor in alleviating learners’ possibly heavy-loaded cognitive capacity. The actual study had put the task familiarity factor into consideration.
Shirai (2007) however raises the difficulty of obligating learners to use past tense in a narrative. She explains that although narratives largely recount past happenings, native speakers sometimes employ historical present in conversational narratives. To further obligate learners’ use of past tense in the narrative task, the present study preset the historical time phrase “A long long time ago...” in Chinese as the opening of each narrative scene and presented it in each of the pre-task Chinese summaries. The pre-task Chinese summaries were read aloud for learners. These were to effectively direct learners towards the historical timeframe. A pre-task Chinese summary was given prior to each test and feedback sessions to orient learners to the story meaning of the assigned scene.

“Archer God’s Shooting Ten Suns and Chang Or’s Flying to Moon” consisted of four different but connected scenes: A) the ten suns caused disasters to the earth and the God of The East asked Archer God to use arrows to scare his ten sun children back to the Palace of Heaven; B) Archer God went down to the earth with his wife and helped The Emperor of the earth get rid of the ten suns; C) Archer God shot the nine sun children dead out of rage and the furious God of The East ousted Archer God and his wife from the Palace of Heaven; D) Archer God went to ask for the Medicine of Long Life from the God of The West to avoid facing the human reality of getting old and dying in the end on earth, but the wife of Archer God flew to the moon alone from wrongly taking the medicine. These picture prompts may relieve some of learners’ cognitive pressure of creating the story meaning from scratch. Learners may thereupon have more cognitive capacity to deploy their declarative knowledge of past tense in their spoken narratives. The cartoon strip of each narrative part is given in appendix IV. Each of the cartoon strips, its prior English and Chinese instructions, and its Chinese pre-task summary (see section 4.8, appendices V and VII) had been
piloted with the researcher's native speaking and non-native speaking personal contacts. Each had elicited the native speakers and non-native speakers' adequate use of past tense.

Same as the pilot study stories, each of the narrative parts consisted of the combination of foreground and background elements. The combination shapes the universal structure of a narrative, as well as elicits learners' potential native use of past tense for both elements. However, the present study did not identify the particular foreground and background elements of the stories, since they were only vehicles to formulate the structure of the narrative task. Moreover, the study's focus was on learners' use of past tense elicited by their narratives, instead of the distribution of past tense in the different structures of the narratives.

4.7.4 Narrative in the Spoken Mode

Narratives in both the spoken and written modes were employed in Bardovi-Harlig's (1992a, b, 1995, 1998) studies of verb morphology in narrative discourse to comprehensively investigate learners' tense interlanguage in the two modes. The present study however limited its investigation to the spoken mode because of its primary focus on the effects of the different oral recast treatments. Bardovi-Harlig (1992a, b, 1995, 1998) compared the results of the spoken and written narratives and found that learners' rate of appropriate past tense usage in the written mode was higher than that of the spoken mode. This suggested that learners may need some intervention to enhance their appropriate use of past tense in spoken narratives. The spoken mode may serve as a vehicle for recasts to perform prospectively their corrective function.
Apart from that, recasts are originated from the spoken mode in mothers’ or caregivers’ recasting children’s utterances in their first language acquisition (Farrar, 1990, 1992). Recasts given in the spoken mode to learners’ spoken narratives would therefore appear more appropriate. Moreover, the written mode, which normally allows more online planning due to its lesser requirement of immediacy than the spoken mode (Bygate, 1987) and thereby facilitates learners’ accuracy and complexity of using the target form (Ellis, 1987; Foster & Skehan, 1996; Skehan & Foster, 1997, 1999), may not permit enough opportunity for recasts to perform their corrective function. The written narrative mode was thus not the instrument option for the present study.

As discussed in chapter 3 section 3.7, the speaking mode can be a potential variable to the present study learners’ attention to past tense and noticing the formal focus in recasts. Some task control measures, for example planning time entailed in the writing mode to accommodate learners’ controlled processing, may be helpful in easing the impact of the speaking mode on learners’ attention to form during their spoken narrative task. Pre-task planning was thus included in the present study.

In the process of collecting spoken data from learners in the actual study, their spoken narratives were first of all recorded to the audio digital tape recorder via the clip-on microphone fixed on each student’s collar. The tape recorder and microphone were tested with each student prior to the commencement of each session. The process of recording is recounted in appendix V. The digital data were then transferred and stored in the computer for the researcher and second transcriber/coder to transcribe and analyse. The transcription rules for transcribing the raw audio data, the coding rules for coding tokens of past tense in learners’ spoken narratives, and the recurring issues arisen in learners’ stimulated recalls, for both the researcher and the
second transcriber/coder's reference, are detailed in appendix VIII. Appendix IX also illustrates some transcription and coding samples to exemplify the rules.

4.7.5 Stimulated Recall Procedures

From sections 4.7.1 to 4.7.4 above, the narrative instrument of the quantitative experimental approach has been identified and justified in terms of its narrative discourse type, narrative structure, task familiarity to learners, and use of the spoken mode and elicited narrative, in sustaining research reliability and data validity. The qualitative approach of stimulated recall in the present study on the other hand made use of question prompts and selected video episodes as instruments, to stimulate learners' recollection of their perceptions of the feedback received at a specific time. The pre-determined question prompts and selection criteria of video episodes are defined in chapter 6, preceding the analyses of students’ stimulated recall data.

4.8 Data Collection Procedures

Sections 4.5 and 4.7 have characterized the participants and the instrument of the present study. This section will proceed to recount the procedures that the participants interacted with the instrument from the beginning to the end of the study’s data collection. Procedures which were different from those of the pilot study will be described to underline the refinement made in view of internal validity. Attention to potential threats to internal validity and corresponding ways implemented to attempt to secure internal validity will be given along the procedures’ account. The extent that the study’s procedures allowed results to be generalized outside the context of the study will also be evaluated. The following will delineate the entire data collection procedures in chronological order.
A month prior to the three-month data collection period from March to May 2007, the target school was approached based on the researcher’s personal contact with the school principal. The principal then referred the researcher to the school’s English Language subject coordinator for all the subsequent arrangements. The coordinator first of all informed the researcher of their approach of teaching English according to students’ different proficiency levels. In practice, the researcher was told that students in Secondary 1-3 (S1-3) were divided into advanced, medium and low streams. Among them, the medium stream comprised the largest number of students with 36 students in each of its three classes. The advanced and low streams constituted only one class of 36 students each. Having the largest population and students’ medium proficiency, the researcher decided to request to the English coordinator about obtaining the three S3 medium-stream classes as the participants of the study. The suitability of this level has already been stated in section 4.5.1.

Two weeks before the start of data collection in March 2007, the S3 English Language subject coordinator (Teacher A), who was also a teacher of one of the three medium-stream classes, arranged a meeting for the researcher together with all the three medium-stream English teachers (Teachers A, B and C). Teacher A took the initiative to invite the researcher to start with his/her class (Class 1). With the help of Teacher A, the researcher went to the class, introduced to the students about the general purpose of the study, and distributed the consent form (see appendix I) to them one week before conducting the study.

Steps employed by the researcher in the first meeting with Class 1, as well as Class 2 and 3, were: 1) greeting the English teacher and the students, and the researcher introducing herself; 2) distributing the greeting / cover letters in Chinese (see appendix I) addressing to each student and his/her parents; 3) telling the students
that the general purpose of the research was to know more about their EFL learning;
4) telling the students generally that they would be asked to speak in English; 5)
telling the students that their individual time and dates meeting with the researcher at
school were stated in the letters addressing to each of them; 6) mentioning that the
location of the meetings would be notified on that day; 7) stressing that audio and
video taping would be conducted during the assigned meetings with the researcher; 8)
clarifying that all data collected would be kept confidential and anonymous, only for
the researcher's PhD studies, and any report, presentation or publication that arose; 9)
distributing the consent forms in Chinese (see appendix I); 10) emphasizing that the
research and all the time and dates had been approved by the school as shown in the
school's chop on the letters and consent forms; and 11) teacher helping with
collecting all the returned consent forms signed by students and their parents within
the week. All these steps were conducted in Cantonese. More detailed procedures are
demonstrated in appendix V.

As expounded in section 4.5, the distribution of the two experimental groups
(corrective recast and normal recast) and one control group (content-only feedback)
was spread across each of the three English medium-stream classes (Class 1, 2 and 3),
to balance the variables of students’ different English medium-stream classes and
teachers and their individual proficiency levels. One uncontrolled variable was having
different students within a class as well as different students from the three different
English classes participating in the study in different time slots and dates. This was, as
mentioned in section 4.5, due to the study’s time period clashing with the students’
spring term (March-May). However, the study tried its best to control such a variable
by spreading the different groups (corrective recast, normal recast and control) across
different timeslots in different days and weeks. Appendix II illustrates the actual time and date arrangements for each participant from each English medium-stream class.

Under the unavoidable variable of students’ different schedules, the researcher had to collaborate with the three English medium-stream classes in different weeks but along the same procedures. As shown in appendix II, Class 1 was the first class participating, followed by Class 2 and 3. The delayed posttests were administered three weeks later, which was the longest period the school term allowed. Other than the students’ school Easter’s holidays when only Class 1 was doing their delayed posttests, there was no break time in the middle in order to finish everything before students’ final exams in June. At the end of each meeting with each student from each class, each student was given a set of vocabulary about food (see appendix X), as a practice of letting them learn something from each meeting. After meeting with all the three classes, the researcher distributed a debriefing letter (see appendix XI) to each participant to explain the research focus in more detail, plus attached a stationery gift as a token of appreciation.

Due to the unchangeable variable of students’ different participation schedules during their hectic school term, students who were assigned to a less packed timeslot did their pretests immediately before their feedback sessions. Students who were assigned to a more packed timeslot did their pretests in one timeslot some time before their feedback sessions in another timeslot within the same school day or on the next immediate day. A certain timeslot being packed or less packed was determined by the final number of participants in a class, the students’ free lesson time allowed by their teachers, or students’ own off-lesson time in their independent study session, lunch time or after-school period. Although some students’ pretests were conducted some time before their feedback sessions, all students’ immediate posttests were arranged
immediately after their feedback sessions to examine the immediate effect of the feedback. The variable of splitting pretest and other sessions for some students was controlled by having equal number of learners from the three feedback groups experience such broken timeline, to reach a balance (see appendix II).

Students with more free time were also assigned to have stimulated recall sessions. As delineated in section 4.5.1, a subset of 35 students were invited to have Cantonese stimulated recall interviews with the researcher in addition to their three tests and one feedback session conducted in English. Only a subset was chosen because of the key problem of time constraint during students’ hectic school term. The stimulated recall interviews were conducted after immediate posttests (see appendix II). This was to allow the video-taped feedback sessions to serve as bases for the stimulated recall interviews, and ensure that the stimulated recall interviews would not raise students’ awareness of using past tense in the immediate posttests if they were conducted beforehand.

The stimulated recall interviews were arranged immediately after the immediate posttests, so that the time gap between each feedback session and stimulated recall interview would not be too wide to corrupt students’ memory (Gass & Mackey, 2000, p.85). Moreover, the stimulated recall interviews were audio-taped in the same manner as the other three tests (pretest, immediate posttest, and delayed posttest). Appendix V illustrates the transcript of the precise steps and instructions in Cantonese (presented in English as well), adopted to lead each student to proceed at ease from the pretest, feedback session and immediate posttest to the stimulated recall interview. The estimated time allocations in each stimulated recall session, including the transition time from the preceding immediate posttest, the instruction briefing
time, the practice time, the video replay time and the recall time or the interaction between the researcher and the participant, are stated in appendices II and V.

Stimulated recall students' corrective recast, normal recast or content-only feedback sessions were all video-taped to prepare for their following interviews, which explored their perceptions towards the feedback they had received and responses they had exhibited. To ease students' potential discomfort with expressing their thoughts at length, the interviews were conducted in their first language, and the researcher presented the interview as simply a sharing session of their thoughts. Learners were also instructed to recall thoughts at the time of the video rather than at the time of the interview.

To stimulate students' recalls, the feedback session videos were replayed. During the process, the researcher paused at video extracts which showed students' certain responses during their narration or after certain feedback was given, and prompted students with a generic question to stimulate their recalls of their original thoughts at that time. Prior to all the actual video pausing and question prompting, students were invited to pause the video, ask questions at certain points, and then restart the video with the remote control. The whole process of pausing the video and asking questions had been modeled to familiarize students with the forthcoming procedures, as well as train them to pause the video before initiating to speak about it (Gass & Mackey, 2000).

For the purpose of sustaining internal validity, all video extracts were selected based on students' apparent responses on the video, to authentically draw students' attention to their specific responses at that time. This aimed at motivating students' ability to articulate rather than fabricate their cognitive experiences at that time. Furthermore, questions prompts conducted in Cantonese were designed to be generic
and orient students to the feedback time with key phrases, "what were you thinking" and "at that time", without asking students to explain anything in-depth about their specific responses. This was to avoid the threat to veridicality in leading students to give answers preferred by the researcher or to create their own answers at the time of the interview.

Moreover, the researcher did not give any specific responses to students’ recalls except backchannelling (e.g. ok, um), to allow students’ free articulation of their stimulated thoughts without interrupting them, accept students’ answers of not remembering anything, and avoid distorting students’ authentic recalls (Gass & Mackey, 2000, p.60). In cases where students gave no responses, gave too lengthy responses, or started to analyse their responses on the video instead of recalling what actually happened, backchannelling was also employed before leading students to other video extracts and recalls. All these were to sustain internal data validity.

Despite students’ different participation schedules, the time allocation, procedures as well as instructions given in each student’s pretest, immediate posttest, and delayed posttest were designed to be similar (see appendix V). This arrangement attempted to standardize data collection from each student to sustain internal validity. The time allocation for each test, feedback and stimulated recall sessions was derived from the researcher’s modeling the entire process of each session with a group of personal non-native speaking contacts and the school term schedule at that time.

The study had spread the distribution of the different experimental and control groups across the three different English medium-stream classes and different timeslots during the school term to sustain internal validity. Apart from that, the present study also applied the practice of having different narrative parts for different tests and feedback sessions, by counterbalancing the order of the four different parts
of the same mythical story. These were to reduce the potential practice effect of students’ remembering materials of the same narrative from one test to another, and avoid participants in the same school communicating with each other about which particular narrative part appeared in a particular test or feedback session.

The practice effect that students may learn from the pretest about the study’s focus would not happen since there was no feedback treatment given to the pretest. Moreover, students’ having learned from the similar narrative format in the pretest and feedback session and then performing well in both the immediate posttest and delayed posttest was controlled by administering the different story parts, with different event happenings, in the different sessions. The preemptive use of stimulated recall interviews (Gass & Mackey, 2000, p.58) based on students’ feedback sessions and their perceptions of the feedbacks may help suggest if they learned from something else other than the feedbacks given.

As specified in section 4.7.3, the traditional Chinese mythical story “Archer God’s Shooting Ten Suns and Chang Or’s Flying to Moon” was divided into four scenes or parts. Parts A, B, C and D involved event happenings in sequential order, but they were not given in sequential order or in a fixed order to any students along their chronological sessions (pretest, feedback, immediate posttest and then delayed posttest). Their order counterbalancing was to avoid one part being too easy or too difficult as material for one test session or for students receiving or having received one type of feedback. Table 2 below shows the different ordering of the four different parts of the mythical story in students’ different test and feedback sessions. Individuals from the different feedback groups were assigned to the different versions of ordering according to their schedule order. For example, the first three students receiving corrective recast, normal recast, and content-only feedback respectively.
would all narrate the first version of narrative ordering; the second version was for the
second batch; the third version for the third batch; the fourth version for the fourth
batch; and the cycle repeated for following batches.

Table 2—Counterbalancing the Different Story Narrative Parts

<table>
<thead>
<tr>
<th>Narrative Order Version</th>
<th>Pre-test</th>
<th>Feedback Session</th>
<th>Immediate Post-test</th>
<th>Delayed Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR1, NR1, CNT1 (1)</td>
<td>B (2)</td>
<td>A (2)</td>
<td>C (2)</td>
<td>D (2)</td>
</tr>
<tr>
<td>CR2, NR2, CNT2</td>
<td>D</td>
<td>B</td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td>CR3, NR3, CNT3</td>
<td>C</td>
<td>D</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>CR4, NR4, CNT4</td>
<td>A</td>
<td>C</td>
<td>D</td>
<td>B</td>
</tr>
</tbody>
</table>

(1) CR (corrective recast), NR (normal recast), and CNT (control) students receiving narrative order version 1 (B, A, C and then D) in their chronological test and feedback sessions.
(2) Cartoon strip parts (see appendix IV): A (First part of the Chinese Myth), B (Second part of the Chinese Myth), C (Third part of the Chinese Myth), D (Fourth part of the Chinese Myth).

As shown in table 2, individuals from the feedback groups were spread across the different ordering, to balance the potential difficulties incurred by the different ordering among all participants receiving different feedbacks. Appendix VI details the near-even distribution of participants’ different feedback types across their different genders, English classes, cartoon-strip presentation orders, and stimulated recall sessions. Data validity was expected to be thereby secured. The narrative order arrangements in table 2 did not involve the sequential order of parts A, B, C and then D. The reason was to eliminate the potential outcome that some students exceedingly benefited from the sequential order presentation of the Chinese mythical story.

The pilot study however adopted the way of fixing a certain cartoon strip to be used in a certain test or feedback session (see appendix III), which bore the threat of having an easier or more difficult one as the instrument of a certain session. The actual study’s choice of mixing the order of cartoon strips, so that every different cartoon strip could be used in any test or feedback session, attempted to reduce the
threat of having the results being affected by the different difficulty levels of the
different cartoon strips.

Instructions (see appendix V), introduction of the main story characters only
in the pretest (see appendix IV), and a Chinese summary (see appendix VII) of each
of the parts were given to each student prior to their narration. With the help of the
prior Chinese summary of each cartoon strip, it was predicted that learners would not
be confused with the non-sequential order of the part being assigned to them in a
specific session. Another purpose of presenting the Chinese summary of each part
was to enhance students’ familiarity with the material, in order to lessen their
potential cognitive load. Students were given pre-task planning time to read the
Chinese summary, while the researcher was reading the summary aloud to capture
their concentration. Then, time was given for them to read the cartoon strip of each
narrative session. Appendix V shows the exact pre-task planning time, as well as the
actual task time, allocated for each narrative session. It was predicted that students
would then have more cognitive space to focus on the use of past tense during their
narratives.

Different from Sangarun’s (2005) study, the present study did not employ any
guided planning. This was because first of all guided planning on the use of irregular
past would be too much a giveaway and defeat the purpose of eliciting learners’
spontaneous use of irregular past and errors for recasts to target. Second of all, the
present study focused on the effectiveness of recasts, not the effectiveness of pre-task
planning. Pre-task planning in the present study was only one of the task control
measures to lessen the resource-depleting task complexity. Learners’ focus on form
during their ongoing speech performance may be aided by recasts, which was the
study’s focus, in addition to the peripheral help by pre-task planning.
The present study’s data collection procedures were developed from refining the pilot study’s procedures. In terms of changes in procedures, the actual study did not provide students with a prior English workshop. This was firstly due to the time constraint of the students’ school term, and secondly the students were having their own English lessons concurrently. To control the level variable in the pilot study, the actual study selected students all from the same medium stream. Moreover, the pilot study’s schedule in the summer was less tight and inconsistent than that of the actual study. The actual study was not conducted in the summer because a large enough population of students could not be expected during non-school term periods.

The practices of 1) spreading the different feedback groups across the different classes, different timeslots and different ordering of the narrative parts, 2) standardizing the time allocation, procedures and instructions of each test, feedback and stimulated recall session, 3) channeling students’ attention and recall to the time of their feedback sessions during the stimulated recall interviews, and 4) providing a prior Chinese summary to familiarize students with the particular non-sequential narrative part arisen from the order counterbalancing all aimed at upholding the internal data validity of the present study. Concerning the study’s external validity, the experimental condition of giving recasts and content-only feedback in isolation may not reflect the effectiveness of the different feedbacks in actual classroom interactions between teachers and students. However, as explained in section 4.7.2, interaction may involve the variables of interlocutors’ meaning ambiguity and misinterpretation. Moreover, interaction may not trigger students’ use of past tense in a systematic or secured manner as much as elicited narratives in a more controlled setting.
4.9 Data Analysis

This section will particularize the different means of analyzing the quantitative data. The means for analyzing the qualitative data will on the other hand be detailed in chapter 6, where stimulated recall findings are analysed. The present study first of all adopted the use of ratios to compare the different students’ rates of attempting past tense relative to their individual lengths of narratives. Students’ rates of attempting past tense were calculated through dividing their number of irregular past tokens by the total number of obligatory contexts for irregular past in their spoken narratives. The counting of irregular past tokens and obligatory irregular past contexts and computation of the ratios were performed by the Microsoft Word’s program called Macros.

Although the use of ratios in comparing students’ patterns of attempting past tense in their different lengths of narratives was also employed by Bardovi-Harlig (1992a, b, 1995, 1998) and Han (2002) and it allows comparison of different variables on the same basis, it makes data less precise. For example, 50% of past tense attempts could be represented by a student’s attempting two past tense tokens within the four obligatory contexts in his/her narrative, or twenty past tense tokens out of forty obligatory contexts. Such practice of ratio also makes the different students’ abilities of producing spoken narratives with past tense ambiguous. The latter case in the previous example definitely performs better than the former one; however, the ratio, 50%, obscures the underlying result.

The trade-off effect of forsaking precise comparison among students’ attempts of past tense from the use of ratios is unavoidable, since students’ different degrees of verbosity are variable. Bayley (1996) adopted the use of setting a cut-off point, for instance only counting a fixed number of verbs first appeared within learners’
different lengths of responses, to try to achieve standardization. However, valuable data that appeared later in learners’ speech would be lost. For the purpose of comparing variable data on the same basis without losing any count of students’ past tense tokens, the present study decided to follow Bardovi-Harlig (1992a, b, 1995, 1998) and Han’s (2002) use of ratios.

In Shirai’s (2007) criticism of Lardiere’s (2003) adherence to the obligatory context analysis, Lardiere’s use of ratios of learners’ morphemes supplied in obligatory contexts (SOC) to measure their mastery level of the morphemes has been condemned as obscuring learners’ accuracy. Shirai notes that if learners overuse a certain morpheme, then the ratio of SOC would be made falsely high. This may then run the risk of missing the full picture of learners’ mastery and incur comparative fallacy. Comparative fallacy is proposed by Bley-Vroman (1983) to unveil the downside of relying on the native obligatory context to analyse learners’ output, which is not respecting the autonomous nature of learners’ language.

Comparative fallacy has also been argued by other researchers as being present in other researchers’ analyses. First of all, Lakshmanan and Selinker (2001) claim that Lardiere’s (1998) early study has committed comparative fallacy by relating a learner’s deficiency to his/her lack of past tense marking in obligatory contexts; not marking past tense according to the native speakers’ grammar in other words. They on the other hand argue that the learner would have shown his/her proficiency if Lardiere (1998) used the aspect and discourse hypotheses to interpret the results; that is using telicity and foregrounding as the indices to closely measure the learner’s interlanguage development. However, Lardiere (2003) later defends her result analysis by the fact that the learner’s past marking was not determined by
telicity and foregrounding at all. The aspect and discourse hypotheses were thereby refuted.

Lardiere (2003) also accuses the aspect hypothesis, which hinges on the semantic aspect of verbs to mark past tense, of committing comparative fallacy by universally assuming learners’ L1 semantic representation of verbs the same as that of the target language. She reckons that this in a way induces another type of obligatory context analysis. However, Shirai (2007) disagrees on equating the aspect hypothesis with the obligatory context analysis, because the former has been targeting the form-meaning correlation between tense marking (past/non-past) and its lexical aspect meaning (perfective/imperfective). This actually addresses all uses of the different morphological forms, independent of the obligatory context of marking both the perfective and imperfective meanings in past tense.

Notwithstanding Shirai’s support of the aspect hypothesis, the present study chose to use the target index of obligatory context analysis to measure learners’ use of past tense. This is because as Lardiere (2003) notes, the aspect hypothesis does not predict the end state of learners’ L2 development, lending no reference to where learners should ultimately aim at. The present study adopted Lardiere’s way of target analysis to probe learners’ generic past tense usage in both the perfective and imperfective aspects as well as the foreground and background narrative elements; however, it did not intend to measure learners’ form accuracy. As specified in section 4.6.1.2, learners’ emergent interlanguage form of past marking, for example through *be* verb-raising in *was take* or *was took* instead of inflection *took*, was counted as past tense tokens in addition to their target marking. The study’s disregard of learners’ interlanguage developmental stage but adoption of learners’ interlanguage form
marking was to realistically and closely capture learners’ interlanguage evidence (Doughty, 2004), without compromising their target usage at the same time.

To avoid the potential comparative fallacy as reminded by Shirai (2007), the present study did not count learners’ overuse of past tense tokens in inappropriate contexts, for example use of past tense when presenting immediate dialogues between characters in the story narrative. Thereby, misrepresentation of learners’ SOC ratios can be kept to the minimum. This may somehow address what Shirai (2007) suggests is the best way to avoid comparative fallacy, which is to take care of learners’ use of form in both the obligatory and non-obligatory contexts.

Concerning the second matter of data analysis, the present study was conducted solely by the researcher. The researcher therefore played the roles of implementing the entire data collection process, transcribing and coding audio data from students’ pretests, feedback sessions, immediate posttests, stimulated recall sessions and delayed posttests. To reduce the threat of researcher bias, a second transcriber for transcribing the audio data; the same second coder for coding the past tense data and stimulated recall data; and the same second translator, who is a Cantonese native speaker and a fluent learner of English, for translating the stimulated recall data from Chinese to English was invited. She was invited to transcribe 12% of the audio data of learners, code 12% of the transcribed past tense data, and translate and code 12% of the transcribed stimulated recall data.

By means of calculating simple agreement rates via dividing the similar tokens by both the similar and different tokens the two parties derived, 87% audio data inter-transcriber agreement, 80% stimulated recall inter-translator agreement, and 83% stimulated recall inter-coder agreement were reached. The agreements in coding the obligatory irregular past contexts and the irregular past forms being tested were
examined statistically by the *intra-class correlation (ICC)* (Shrout & Fleiss, 1979).
The rationale, limitation and results of the ICC are shown in chapter 5. A statistical
tool was used to estimate the inter-rater reliability in assessing both the obligatory and
learners’ irregular past tokens, because the tokens were the foundation of assessing
learners’ quantitative development and gaining a more objective reference of the
inter-rater reliability is especially crucial. The transcription and coding rules (see
appendix VIII) were given and briefed to the second transcriber, coder or translator
during the training or briefing session.

The coding of participants’ use of irregular past includes their clear use of
irregular past verbs as well as their inaccurate or developing use of irregular past (e.g.
was took). The coding of obligatory contexts for irregular past includes participants’
non-past tense use of irregular verbs in addition to the above two categories. Regular
past, verbs with no change of vowel from their base form to past form (e.g. have/had,
make/made), model verbs, verbs with the same past form as their base form (e.g. cut,
read), irregular past forms in non-finite contexts, repetitions, unclear irregular past
forms, and irregular past forms in participants’ presentation of dialogues between
characters in the narrative were not counted in the analyses. Appendix VIII details the
coding rules of irregular past tokens and obligatory contexts for irregular past. Lay
language in brackets had been provided (see appendix VIII) to simplify technical
terms in the transcription and coding rules, for the second transcriber/coder’s
reference. Several samples of the three test data transcription and coding, the
feedback session transcription and coding, the stimulated recall interview Cantonese
transcription and English translation are illustrated in appendix IX. Chapter 6
illustrates how exactly the stimulated recall data were coded.
The researcher first selected valid stimulated recall data and identified thematic issues (see chapter 6 and appendix VIII) that recurred throughout the entire valid dataset. After that, the researcher invited the second coder, who also helped translate the Cantonese speech to English, to code 12% of the valid stimulated recall dataset and check whether the recurring thematic issues (see chapter 6 and appendix VIII) identified corresponded legitimately to the stimulated recall segments transcribed and selected previously.

Regarding other statistical tools used, the *repeated ANOVA* was first of all performed to examine the overall interaction effect between the group and time factors. Moreover, the quantitative data of the different feedback groups’ pretests, immediate posttests and delayed posttests were statistically tested by the *one-way ANOVA*, for specifically comparing and showing the significant differences among the different groups at different time points. The *post hoc test* comparison was used to further dissect which feedback types contributed to the significant differences.

ANOVA statistically tests whether or not the means of several groups are all equal, and extends the practice of *t-test* to testing the means of more than two groups. Doing multiple *t*-tests to test means of more than two groups would incur a *type I error*. This is why ANOVAs are rather secure and useful in comparing two, three or more means.

Type I error is the error of wrongly rejecting a true *null hypothesis* ($H_0$). Therefore, ANOVA was also used to compare the stimulated recall group and non-stimulated recall group, to eliminate the possibility that the implementation of stimulated recall may affect the delayed posttest performance of the stimulated recall group.

Apart from data relevant to gauge learners’ use of past tense across time after receiving different feedbacks, the *two-way mixed ANOVA* was also used for data beyond measurement of learners’ performance. The two-way mixed ANOVA was
used to show the main time and group effects and their interaction effects of the 9 types of formal tokens, both past and non-past forms, used by learners across time and feedback groups. Lastly, the Mann-Whitney U test was used to look into how different the two feedbacks (normal and corrective recasts) were performed in the data collection based on the specified variables. The use of this non-parametric test was because the data were not in normal distribution. Display of data obtained from these tests is provided and analysed in chapter 5. The rationales behind using the tests will be elaborated in chapter 5.

4.10 Ethics

Reflecting on the study’s ethical issues, procedural steps had been conducted across the data collection period to inform the students of the research ethically, without revealing too much to them to affect data validity later. From undertaking the study, ethics were helpful in constraining the research to put the knowing right of participants into consideration. Participants, as humans with individual rights and independent thinking, were considered reserving the right to know about the role of the researcher and the relevant investigation, and the right to reject the researcher’s invitation. A Chinese cover letter (see appendix I) briefing the purpose of the research and a Chinese consent form (see appendix I) asking for permission from the students and their parents to participate in the research were therefore given prior to the data collection.

However, to balance ethics and research validity, the study only told participants about the general area of investigation, which was their EFL learning. This was to prevent students from anticipating what they should do to fulfil the research purpose, and thereby distorting the results. As in the pilot study, the final
debriefing stage aimed at unfolding the genuine research purpose to compensate the previous lack of transparency. Overall, ethics are reminders for researchers to avoid perfecting their research at the expense of participants’ rights.

4.11 Conclusion

Robinson’s (2005) attentional model and view of task complexity have been shown more directional, educational, comprehensive and specific than Skehan’s in chapter 3. They therefore lent support to the present study’s hypotheses and task control. The present study largely adopted Robinson’s 1) task complexity position to control its task design to maximize learners’ attention to the form and recasts; 2) position that task complexity requiring the use of rather novel target forms can facilitate learners’ learning of the forms as well as their tendency to seek help from recasts; 3) and specific measure of learners’ use of the target form to look into their performance of past tense in particular.

Relevant research (Robinson, 2001a, 2007; Skehan, 1998; Skehan & Foster, 2001) probing into the relation between task complexity manipulation and learners’ formal accuracy and complexity may inspire the present study with how to best control the potential variable of task effect. It is hypothesized that the task variable will obstruct learners’ noticing recasts and hence the effectiveness on restructuring Hong Kong EFL learners’ use of past tense. The present study predicated its task demand control on Robinson’s (2001a, 2005; Robinson & Gilabert, 2007) more linguistic-learning-favouring task framework, alongside Skehan’s (1998; Skehan and Foster, 2001) more procedural-control-focused task framework, of how task complexity influences learners’ attention to and use of past tense.
The resource-directing dimension proposes that tasks eliciting learners' use of past tense may push learners to direct attentional resources specifically to past tense, and realise new linguistic encoding of meaning in speech production (Robinson, 2005). Based on such premise, the present study made use of a mythical narrative task with a Chinese historical background. It was expected to elicit learners' formal complexity in attempting past tense and allow recasts to target. Also using narrative as the task, Robinson's (1995) early study made use of the presence or absence of physical contextual support to elicit learners' past tense. Revesz (2009) however points out that it is problematic to use physical contextual support to perform the here-and-now variable; because it would be difficult to disentangle the confusion between effects of linguistic difficulty and picture prompt availability on task performance.

As a result, the present study adopted the use of the mythical story's historical timeframe to elicit learners' use of past tense. It thereby separated the there-and-then variable from the use of cartoon strip picture prompts. Picture prompts were used specifically to lessen the performative/procedural demand on learners' cognitive processing (Ellis, 2003). The storyline was tightly structured in serial and ordered pictures to avoid any loose ideas diverting learners' attention to form (Skehan, 1998; Skehan & Foster, 2001; Ellis, 2003; Tavakoli & Foster, 2008, 2011). Tavakoli and Foster (2008, 2011) found in their study that narratives structured tightly from the starting picture prompt to the last one fostered learners' formal accuracy, because of their cognitive burden released from organizing the content development. They also demonstrated that storylines with only foreground moving events forward along the timeframe were less complex, which promoted formal accuracy, than one with background as well. According to Robinson's (2001b) map-marking task complexity
example and Ellis’s (2003) single/dual task demand condition, the presence of picture prompts may by themselves avoid imposing the dual task demands of creative meaning construction and language expression on learners.

According to Robinson’s resource-depleting task complexity and Skehan’s limited capacity theory, tasks with less planning time (Ellis, 1987, 2005; Foster & Skehan, 1996; Robinson, 2001a, 2001b, 2005; Robinson & Gilabert, 2007; Skehan & Foster, 1997), loose task structure (Skehan & Foster, 1997, 1999; Robinson, 2001a, 2001b, 2005; Robinson & Gilabert, 2007; Tavakoli & Foster, 2008, 2011; Tavakoli, 2009) leading to dual tasks, and low familiarity (Robinson, 2001a, 2001b, 2005; Robinson & Gilabert, 2007; Skehan, 1998; Skehan & Foster, 2001) would affect learners’ formal attention. Robinson’s task complexity notion along the resource-directing and resource-depleting dimensions can be seen as both synergistic in promoting real-world performance requiring new as well as existing linguistic forms, and conflicting with the latter countering the former. Therefore, besides structuring the storyline with fixed picture prompts to ease learners’ new form-meaning mapping in speech production, the present study also familiarized learners with the story meaning through summaries in their L1 and cartoon strip reading time before task performance. These were to avoid making the task resource-depleting, by enhancing learners’ prior knowledge of the story meaning and planning for the use of past tense.

Borrowing Ellis’s (2003) itemized breakdown of task complexity, criteria chosen to ease learners’ procedural pressure in the present study involved: 1) use of pictorial input medium, 2) dynamic information type with substantial characters’ actions instead of abstract, 3) a small amount of information with relationships among a few characters instead of many, 4) familiar information, 5) closed scope with fixed story meaning instead of open meaning, and 6) use of less detail-oriented discourse
mode, for example narratives instead of instructions or arguments. Although Ellis regards that a two-way and dialogic interactive task would be easier, the present study employed a one-way and monologic task. Robinson (2001b) argues that monologic narrative tasks would enhance formal accuracy and complexity along the resource-directing dimension, and formal complexity is precisely the present study’s focus; whereas interactive tasks entailing negotiation of meaning would mitigate learners’ complex utterances, with turn-taking inducing elliptical answers.

All in all, this chapter recounts how the entire study has been shaped and conducted via establishing the design framework, understanding the context of the target school in Hong Kong, selecting appropriate participants and distributing them evenly in different arrangements, identifying the suitable target form, choosing and implementing the instruments of spoken narratives and stimulated recall procedures in ways that best cater for the study’s purpose, executing data collection in structured procedures and instructions, and analyzing the quantitative and qualitative data with relevant statistical tools and recurring thematic categories respectively. All the above actions endeavoured to foster research replication reliability and internal data validity.

4.12 Research Hypotheses

From the research questions posed in the last chapter and the specifications of the present study provided in this chapter, six hypotheses are formulated as follows:

1) Consistent normal recasts are facilitative in promoting Hong Kong Form Three learners’ use of irregular past in spoken narratives in the short run.
2) Consistent normal recasts are facilitative in promoting Hong Kong Form Three learners’ use of irregular past in spoken narratives in the long run.
3) Consistent corrective recasts are facilitative in promoting Hong Kong Form Three learners’ use of irregular past in spoken narratives in the short run.

4) Consistent corrective recasts are facilitative in promoting Hong Kong Form Three learners’ use of irregular past in spoken narratives in the long run.

5) Consistent corrective recasts are more facilitative than consistent normal recasts in promoting Hong Kong Form Three learners’ use of irregular past in spoken narratives in the short run.

6) Consistent corrective recasts are more facilitative than consistent normal recasts in promoting Hong Kong Form Three learners’ use of irregular past in spoken narratives in the long run.
Chapter 5 Quantitative Data Analysis

5.1 Introduction

The thesis proceeds to begin its quantitative and qualitative data analyses for the present study. This chapter first of all displays and analyses the quantitative results of the present study, to statistically test the six hypotheses via the raw results obtained from the experimental approach characterized in chapter 4.

The following parts will analyse the study’s results by the means of a two-way mixed ANOVA, a one-way ANOVA and pair-wise comparisons. They serve different purposes and provide different implications. Moreover, a two-way independent ANOVA was performed to demonstrate that the stimulated recall interviews did not cause any interfering impact to the stimulated recall subset’s delayed post-test results held three weeks afterwards. Some data analyses beyond the research questions will also be presented to provide insights into the study’s results. The quantitative result analysis provided in this chapter will be followed by the qualitative result analysis of the present study’s stimulated recall data, to suggest further explanations for the statistical findings attained.

5.2 Statistical Analyses of the Study’s Results

The two-way mixed ANOVA with repeated measures at different time points (i.e. each level of test) was employed to investigate how the two independent variables (group and time) affected the outcome variable which was measured by the ratio of past tense used. The ‘group’ variable was a between-subject variable that classified all students into three groups: control (CNT), normal recast (NR) and corrective recast (CR). The ‘time’ variable was a within-subject variable that enabled the investigation of the time effect assessed at three levels - pretest, immediate
posttest and delayed posttest on the treatment. Considering the research design and the nature of the variables, a two-way mixed ANOVA was applied to investigate the significance of the main effects (group and time) and the interaction effect on the outcome measure. In addition, pair-wise comparisons among different treatment groups at different time points were also performed by one-way ANOVA and post hoc multiple comparisons. The test Scheffe or Dunnett T3 was selected as the post hoc test depending on the results of the Levene test of homogeneity of variances. All effects were presented with the effect sizes measured by partial eta square, $\eta^2$, produced by the statistical software. The partial eta square can be defined as the ratio of variance accounted for by an effect and that effect plus its associated error variance within an ANOVA study.

All the statistical analyses were performed by the statistical software SPSS 15.0, and the significance level is predetermined at 0.05. The significance indicators are presented by the exact p-values, except when the p-values are too small (e.g. the p-values of the time and interaction effects in table 3, with exp^8 and exp^20) to have the exact value reported. Normally when the p-values are too small, only the upper limit using the inequality shown is reported.

5.3 Results of the Two-way Mixed ANOVA

To present a global picture of the different effects of the different feedback types (corrective recast, normal recast and content-only) on participants across the different time points (pretest, immediate posttest and delayed posttest), the present study's quantitative results have been presented in figure 1 and table 3 below.

Figure 1 clearly shows that the content-only feedback control group did not seem to change its performance as much as the other two experimental groups.
(corrective recast and normal recast) across the two different time periods (from pretest to immediate posttest, and from immediate posttest to delayed posttest).

Moreover, obvious differences between the two experimental groups and the control group can be shown across time. There however appears a small difference between the two experimental groups’ performance from the beginning to the end. Both the two experimental groups experienced a sharp increase in using past tense in their spoken narratives from the pretest to the immediate post, but a slight decline from the immediate posttest to the delayed posttest.

![Past Tense Used Ratio in the three tests by groups](image)

Figure 1 — Ratios of the Past Tense Used in the Three Tests by Groups
Table 3 – Results of Two-way Mixed ANOVA on the Outcome Measure ‘Past Tense Used Ratio’

<table>
<thead>
<tr>
<th>Group</th>
<th>Total (n=89)</th>
<th>Control (C) (n = 31)</th>
<th>Normal recasts (NR) (n = 29)</th>
<th>Corrective recasts (CR) (n = 29)</th>
<th>Group effect (Greenhouse-Geisser)</th>
<th>Time effect (Greenhouse-Geisser)</th>
<th>Interaction effect (Group*Time) (Greenhouse-Geisser)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>F(2,86) = 3.884, p = 0.024**</td>
<td>F(1.826,157.07) = 62.379, p &lt; 0.001**</td>
<td>F (3.653,157.07) = 11.525, p &lt; 0.001**</td>
</tr>
<tr>
<td>Pretest</td>
<td>0.253 0.170</td>
<td>0.301 0.162</td>
<td>0.207 0.162</td>
<td>0.246 0.177</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate Posttest</td>
<td>0.481 0.213</td>
<td>0.342 0.194</td>
<td>0.533 0.193</td>
<td>0.577 0.177</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed Posttest</td>
<td>0.411 0.230</td>
<td>0.318 0.199</td>
<td>0.462 0.223</td>
<td>0.458 0.244</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < 0.01; \( \eta^2 \) denotes the estimate of partial eta square as a measure of effect size
The results of the two-way mixed ANOVA are presented in table 3, together with the descriptive statistics on the ratio of past tense used in the three tests of the three treatment groups.

Prior to conducting the two-way mixed ANOVA, the Mauchly’s test of sphericity showed that the sphericity assumption did not hold (p=0.014), therefore the Greenhouse-Geisser correction factor was applied to produce a valid F-ratio for the ANOVA. The results of the two-way mixed ANOVA revealed that the effects of the between-subject (group) variable, within-subjects (time) variable and their interaction were all statistically significant at 0.05 level of significance.

When the mixed ANOVA is performed, the sphericity assumption is needed to be checked before interpreting the figures obtained from the model. If the assumption holds, one set of statistics (e.g. F-value, p-value etc) is applied for interpretation; if it does not hold, like the current situation, another set of statistics (i.e. the set with Greenhouse-Geisser correction factor) is applied. The assumption of sphericity is similar to the assumption of homogeneity of variance in ANOVA. In general, the factorial ANOVA assumes the variance of each condition (group) is homogeneous. Sphericity is however a kind of less restrictive compound symmetry. Compound symmetry stays true when both the variances across conditions are equal and the covariances between pairs of conditions are equal. In other words, the variation within experimental conditions is rather similar and no two conditions are any more related than any other two. Sphericity refers to the equality of variances of the differences between treatment levels—the variance between different tests/time points. Therefore, if each pair of treatment levels is to be taken and the differences between each pair of scores are to be calculated, then these differences must have approximately equal variances.
To measure how different the data are from the assumption of sphericity, the Mauchly’s test was applied. Statistical significance (p<0.05) indicates evidence of departure. In case the assumption of sphericity is violated, the F-ratio (a statistics) computed is no longer valid, so adjustment is necessary. Greenhouse-Geisser correction is a kind. It makes correction to the degree of freedom (df), so as to validate the F-ratio statistics and the corresponding p-value.

Specifically, there was significant main between-subject ‘group’ effect on the ratio of past tense used, F (2, 86) = 3.884, p = 0.024, with effect size (η² = 0.083). The results showed that when the time variable was collapsed, significant difference of the mean ratios among the three groups was found. In addition, the within-subject ‘time’ effect on the ratio of past tense used was also found to be statistically significant, F(1,826, 157.07) = 62.379, p < 0.001, with effect size (η² = 0.420). The results showed that when the group variable was collapsed, significant difference of the mean ratios among the three time points exists. Finally, the interaction effect was also shown to be statistically significant, F(3.653, 157.07) = 11.525, p < 0.001, with effect size (η² = 0.211). The results showed that the rates of change of the three groups from the pretest to immediate posttest and to delayed posttest were different.

Referring to figure 1, it can be observed that the mean ratio of the control group did not change much from the pretest to immediate posttest and to delayed posttest. However, the rates of change of the other two groups were much different. The mean ratios of the normal recast and corrective recast groups had improved a great deal from the pretest to immediate posttest, and then followed by a slight decline from the immediate posttest to delayed posttest. It is also noteworthy that the rates of change between the normal recast and corrective recast groups were quite similar relative to that of the control group. In order to investigate the mean ratio differences
among the three groups at each time point, one-way ANOVA and post hoc multiple comparisons were performed. The results are presented in the next section.

5.4 Results of the One-way ANOVA and Post Hoc Test Comparison

First of all, the one-way ANOVA test was applied to detect any statistically significant differences among the three groups, the corrective recast, normal recast and the content-only feedback control groups, started at each time point. Table 4 below shows that the one-way ANOVA exhibited no significant difference among the three groups at the pretest level (F(2,86)=2.414, p = 0.096). This suggests that the three groups participated in the study at more or less the same level, with no group superior to other ones in the beginning. The potential variable of different levels seems to have been controlled by the present study’s method of recruiting all participants from the same S3 (Secondary Three) and medium English level at the same target school in Hong Kong. This also appears to realize the cluster randomization sampling design, in which all the classes are deemed to be equivalent in nature, with no randomization at the individual level.
Table 4 - Results of One-way ANOVA and Post Hoc Multiple Comparisons among the Three Treatment Groups on ‘Past Tense Used Ratio’ at Each Time Point

| Test          | One-way ANOVA | Post hoc Multiple comparison |  |  |  |  |
|---------------|---------------|-------------------------------|  |  |  |  |
|               | Group         | Control vs Normal recast       | Control vs Corrective recast | Normal recast vs Corrective recast |  |  |
|               |               | Mean difference | Std. Error | p   | Mean difference | Std. Error | p   | Mean difference | Std. Error | p   |
|               | Control       | 2.143 0.096 | 0.094 0.043 | 0.099 | 0.056 0.043 | 0.439 | -0.038 0.044 | 0.682 |
|               | Immediate     | 13.371 <0.001** | -0.191 0.049 | 0.001** | -0.236 0.049 | <0.001** | -0.044 0.050 | 0.672 |
|               | Delayed       | 4.148 0.019* | -0.145 0.057 | 0.047* | -0.140 0.057 | 0.056 | 0.004 0.058 | 0.998 |

*p < 0.05; **p < 0.01
Apart from affirming that all the three groups started at the same level before the feedback treatment sessions, the one-way ANOVA was also employed to reveal if the feedback sessions produced any overall significant differences on participants' performance in the immediate posttests and delayed posttests. Table 4 illustrates that the three different types of feedback treatment generated significant differences on participants' past tense usage in their spoken narratives in the immediate posttests \( F(2,86) = 13.371, p < 0.001 \) and delayed posttests \( F(2,86) = 4.148, p = 0.019 \). To further examine which particular feedback type contributed to the significant difference in each of the posttests, the Scheffe or Dunnett T3 post hoc test was performed to compare the past tense usage of the different feedback groups pair-wise.

Table 4 demonstrates the post hoc results, listing that the overall significant difference \( F(2,86) = 13.371, p < 0.001 \) shown in the immediate posttest was attributed to the significant difference between the corrective recast group and the control group \( p < 0.001 \), and also between the normal recast group and the control group \( p = 0.001 \). The difference between the corrective recast and normal recast groups \( p = 0.672 \) was shown insignificant to contribute any statistical significance to participants in the immediate posttest. Regarding the delayed posttest, the overall significant difference \( F(2,86) = 4.148, p = 0.019 \) was illustrated to be caused solely by the significant difference between the normal recast group and the control group \( p = 0.047 \). The differences between the corrective recast and control groups \( p = 0.056 \) and between the corrective recast and normal recast groups \( p = 0.998 \) did not reach the specific statistically significant level.

Two options of post hoc test were given to measure the different group effects, Scheffe or Dunnett T3, to prepare for the flexibility of the results obtained from the Levene test for equity of variance. The choice of multiple comparison post hoc tests
was made based on the results of the Levene test for equity of variance. If the Levene test shows that the variances among groups can be assumed equal, the Scheffe test will be chosen as the post hoc test for multiple comparisons; whereas if the Levene test shows that the variances among groups cannot be assumed equal, the Dunnett T3 will be chosen instead. The Scheffe test was selected as the post hoc test (when equal variance assumed) for multiple comparisons, as it is the most stringent test among all when the sample size of each group is not identical. Regarding the choice of Dunnett T3, the rationale is that it keeps very tight Type I error control relative to other post hoc tests.

5.5 ANOVA for Stimulated Recall and Non-stimulated Recall Groups

Besides implementing the two-way mixed ANOVA, one-way ANOVA test and post hoc test comparison to examine the effects of the different independent variables (groups and time) on the dependent variable (past tense used ratio), a two-way independent ANOVA was performed to help the study exclude one extraneous variable potentially contaminating its delayed posttest data. Since the present study only recruited a subset (39% or N = 35) of the total number of learners (N = 89) to participate in the stimulated recall interview sessions immediately after learners’ immediate posttests, there exists a possibility that the stimulated recall interviews may have raised the subset’s awareness of using past tense in their delayed posttests held three weeks later. This may have affected the validity of learners’ delayed posttest performance. A two-way independent ANOVA was applied to attempt to confirm that the stimulated recall interviews did not cause any interfering effect to the subset’s delayed posttest results. Tables 5a and 5b below show the results of the ANOVA. The
effects of the stimulated recall and the non-stimulated recall groups (student group effect) on the immediate postest ($p=0.976$) and delayed posttest ($p = 0.077$) results were not statistically significant. The insignificant $p$ value demonstrates that the stimulated recall sessions did not cause any significant effect to the subset’s delayed posttest performance.
### Table 5a– Results of Two-way Independent ANOVA of ‘Past Tense Used Ratio’ in Immediate Posttest

<table>
<thead>
<tr>
<th>Treatment group</th>
<th>Total (n=89)</th>
<th>Control (C) (n = 31)</th>
<th>Normal recasts (NR) (n = 29)</th>
<th>Corrective recasts (CR) (n = 29)</th>
<th>Treatment group effect</th>
<th>Student group effect</th>
<th>Interaction effect (Student group*treatment group)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-SR group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 54)</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.480</td>
<td>0.204</td>
<td>0.358</td>
<td>0.160</td>
<td>0.523</td>
<td>0.201</td>
<td>0.570</td>
</tr>
<tr>
<td></td>
<td>F(2,83) = 13.204</td>
<td>F (1,83) = 0.001**</td>
<td>η² = 0.241</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 35)</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.483</td>
<td>0.230</td>
<td>0.316</td>
<td>0.245</td>
<td>0.550</td>
<td>0.188</td>
<td>0.588</td>
</tr>
<tr>
<td></td>
<td>F(2,83) = 0.241</td>
<td>F (1,83) = 0.976</td>
<td>η² = 0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
* p < 0.05; ** p < 0.01; η² denotes the estimate of partial eta square as a measure of effect size

### Table 5b– Results of Two-way Independent ANOVA of ‘Past Tense Used Ratio’ in Delayed Posttest

<table>
<thead>
<tr>
<th>Treatment group</th>
<th>Total (n=89)</th>
<th>Control (C) (n = 31)</th>
<th>Normal recasts (NR) (n = 29)</th>
<th>Corrective recasts (CR) (n = 29)</th>
<th>Treatment group effect</th>
<th>Student group effect</th>
<th>Interaction effect (Student group*treatment group)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-SR group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 54)</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.376</td>
<td>0.224</td>
<td>0.292</td>
<td>0.142</td>
<td>0.425</td>
<td>0.239</td>
<td>0.419</td>
</tr>
<tr>
<td></td>
<td>F(2,83) = 4.097</td>
<td>F (1,83) = 0.020*</td>
<td>η² = 0.090</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SR group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 35)</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.464</td>
<td>0.233</td>
<td>0.359</td>
<td>0.268</td>
<td>0.523</td>
<td>0.188</td>
<td>0.514</td>
</tr>
<tr>
<td></td>
<td>F(2,83) = 0.037</td>
<td>F (1,83) = 0.959</td>
<td>η² = 0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
* p < 0.05; ** p < 0.01; η² denotes the estimate of partial eta square as a measure of effect size
5.6 Measure of Inter-rater Reliability

In order to evaluate the reliability of the principal rater in assessing the number of obligatory contexts for past tense and the number of past tense tokens, a second rater was invited to assess the two measures from the data of 12% of the total participants. The intra-class correlation (ICC) (Shrout & Fleiss, 1979) was employed as a statistical index to measure the inter-rater reliability and the ICC model (2,1) was used. Regarding the interpretation, the ICC values range from 0.00 to 1.00, where 1.00 represents strong reliability among the raters. The results of the ICC on the number of obligatory contexts for past tense and the number of past tense tokens were respectively 0.96 and 0.94 between the two raters. Based on the interpretation guidelines provided by Portney and Watkins (2009), the inter-rater reliability on the two assessments was excellent. The ICC is a general measure of reliability which is based on the variance proportion among within raters and between raters. It is able to provide reliability estimate for both single rating and mean of several ratings. The ICC is appropriate for interval scale, ordinal scale and nominal scale. This is unlike Kappa, which is appropriate for nominal scale only.

5.7 Lists of Verb Tokens

In this section and the following two, descriptive data beyond the research questions of the present study are presented, to provide a broad picture of the verbs that participants used in the different test and feedback sessions. Moreover, the different variables occurred in the normal recast and corrective recast feedback sessions are demonstrated, to show how they differed across the different feedback types. This section first of all reveals the different verbs that appeared in students’ pretests, feedback sessions, immediate posttests, and delayed posttests. The present
study did not measure students’ increased or decreased use of the different verbs across time and their different feedback sessions, because no particular verbs were targeted consistently throughout. However, the showcase of the verbs may provide an insight into both students’ accurate and developing use of the different form of verbs. The different verb patterns are categorized according to the first auxiliary verbs or main verbs shared by the different verb tokens. The different verb patterns have been grouped into 21 categories, for an overview of the patterns. Table 6 shows what the 21 categories consist of.

<table>
<thead>
<tr>
<th>Table 6 — Categories of Verb Patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grouping 1 – verb tokens with no auxiliary</td>
</tr>
<tr>
<td>Grouping 2 – verb tokens start with 'are', 'aren't', or 'are not'</td>
</tr>
<tr>
<td>Grouping 3 – verb tokens start with 'can', 'can't', or 'cannot'</td>
</tr>
<tr>
<td>Grouping 4 – verb tokens start with 'could', 'couldn't', or 'could not'</td>
</tr>
<tr>
<td>Grouping 5 – verb tokens start with 'did', 'didn't', or 'did not'</td>
</tr>
<tr>
<td>Grouping 6 – verb tokens start with 'do', 'don't', or 'do not'</td>
</tr>
<tr>
<td>Grouping 7 – verb tokens start with 'had', 'hadn't', or 'had not'</td>
</tr>
<tr>
<td>Grouping 8 – verb tokens start with 'has', 'hasn't', or 'has not'</td>
</tr>
<tr>
<td>Grouping 9 – verb tokens start with 'is', 'isn't', or 'is not'</td>
</tr>
<tr>
<td>Grouping 10 – verb tokens start with 'make'</td>
</tr>
<tr>
<td>Grouping 11 – verb tokens start with 'may' or 'may not'</td>
</tr>
<tr>
<td>Grouping 12 – verb tokens start with 'must', 'mustn't', or 'must not'</td>
</tr>
<tr>
<td>Grouping 13 – verb tokens start with 'see'</td>
</tr>
<tr>
<td>Grouping 14 – verb tokens start with 'should', 'shouldn't', or 'should not'</td>
</tr>
<tr>
<td>Grouping 15 – verb tokens start with 'will', 'won't', or 'will not'</td>
</tr>
<tr>
<td>Grouping 16 – verb tokens start with 'would', 'wouldn't', or 'would not'</td>
</tr>
<tr>
<td>Grouping 17 – verb tokens start with 'were', 'weren't', or 'were not'</td>
</tr>
<tr>
<td>Grouping 18 – verb tokens start with 'does', 'doesn't', or 'does not'</td>
</tr>
<tr>
<td>Grouping 19 – verb tokens start with 'have', 'haven't', or 'have not'</td>
</tr>
<tr>
<td>Grouping 20 – verb tokens start with 'was', 'wasn't', or 'was not'</td>
</tr>
<tr>
<td>Grouping 21 – verb tokens start with 'saw'</td>
</tr>
</tbody>
</table>

The following 21 tables (tables 7-27) present the individual verb tokens (Grouping 1-21) used by the participants in their spoken narratives in the different test and feedback sessions. Their frequency distributions across the different test and
feedback sessions are illustrated in tables 7-27 in appendix XII. From the leftmost column to the rightmost column in each table in appendix XII, there are the list of individual verb tokens used by the participants, the total frequency count (N) and percentage (N%) of verbs used by all the participants in the different test and feedback sessions, the frequency count (N) and percentage (N%) of verbs used by the control group, the normal recast group, and the corrective recast group in their respective test and feedback sessions.

<p>| Table 7—Grouping 1: List of Verb Tokens with No Auxiliary |
|-----------------|-----------------|-----------------|
| accept          | hear            | pray            |
| affect          | heard           | prepare         |
| agree           | held            | present         |
| answer          | help            | promise         |
| appear          | hold            | punish          |
| arrive          | hope            | raise           |
| ask             | increase        | ran             |
| ate             | invite          | realize         |
| attack          | jump            | refuse          |
| became          | keep            | rely            |
| become          | kick            | remain          |
| beg             | kill            | remind          |
| began           | knew            | reply           |
| begin           | know            | report          |
| believe         | land            | request         |
| bought          | laugh           | return          |
| break           | lead            | ride            |
| bring           | leave           | run             |
| broke           | left            | sang            |
| brought         | like            | save            |
| buy             | listen          | saw             |
| call            | live            | saw...go        |
| came            | look            | saw...say       |
| care            | lose            | say             |
| carry           | lost            | scare           |
| cause           | love            | see             |
| celebrate       | make            | see...fly       |
| change          | make...absorb   | see...happen    |
| cheat           | make...appear   | see...is        |
| clap            | make...are      | see...make      |
| climb           | make...became   | see...were      |
| come            | make...become   | send            |
| cry             | make...can grew | shine           |</p>
<table>
<thead>
<tr>
<th>Action</th>
<th>Verb Phrase</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>damage</td>
<td>make...can plant</td>
<td>shone</td>
</tr>
<tr>
<td>dead</td>
<td>make...can’t farm</td>
<td>shoot</td>
</tr>
<tr>
<td>decide</td>
<td>make...can’t grow</td>
<td>shot</td>
</tr>
<tr>
<td>declare</td>
<td>make...cannot farm</td>
<td>shout</td>
</tr>
<tr>
<td>demand</td>
<td>make...cannot grow</td>
<td>show</td>
</tr>
<tr>
<td>did</td>
<td>make...cannot live</td>
<td>sing</td>
</tr>
<tr>
<td>die</td>
<td>make...cannot plant</td>
<td>sold</td>
</tr>
<tr>
<td>discover</td>
<td>make...did not feel</td>
<td>speak</td>
</tr>
<tr>
<td>discuss</td>
<td>make...did not had</td>
<td>spoke</td>
</tr>
<tr>
<td>do</td>
<td>make...did not have</td>
<td>stand</td>
</tr>
<tr>
<td>drive</td>
<td>make...did not plant</td>
<td>start</td>
</tr>
<tr>
<td>drove</td>
<td>make...did suffer</td>
<td>stay</td>
</tr>
<tr>
<td>dry</td>
<td>make...didn’t have</td>
<td>stole</td>
</tr>
<tr>
<td>eat</td>
<td>make...die</td>
<td>stood</td>
</tr>
<tr>
<td>emit</td>
<td>make...farm</td>
<td>stop</td>
</tr>
<tr>
<td>enjoy</td>
<td>make...feel</td>
<td>suffer</td>
</tr>
<tr>
<td>evaporate</td>
<td>make...feel, felt</td>
<td>suggest</td>
</tr>
<tr>
<td>exit</td>
<td>make...felt</td>
<td>take</td>
</tr>
<tr>
<td>fail</td>
<td>make...get</td>
<td>talk</td>
</tr>
<tr>
<td>fall</td>
<td>make...go</td>
<td>tell</td>
</tr>
<tr>
<td>feel</td>
<td>make...got</td>
<td>thank</td>
</tr>
<tr>
<td>felt</td>
<td>make...have</td>
<td>think</td>
</tr>
<tr>
<td>fight</td>
<td>make...lost</td>
<td>thought</td>
</tr>
<tr>
<td>find</td>
<td>make...think</td>
<td>told</td>
</tr>
<tr>
<td>flew</td>
<td>make...want</td>
<td>took</td>
</tr>
<tr>
<td>fly</td>
<td>make...was</td>
<td>travel</td>
</tr>
<tr>
<td>follow</td>
<td>make...was die</td>
<td>treat</td>
</tr>
<tr>
<td>force</td>
<td>make...was die</td>
<td>try</td>
</tr>
<tr>
<td>forget</td>
<td>make...went</td>
<td>turn</td>
</tr>
<tr>
<td>forgot</td>
<td>make...were</td>
<td>understand</td>
</tr>
<tr>
<td>found</td>
<td>make...were die</td>
<td>use</td>
</tr>
<tr>
<td>gave</td>
<td>make...were living</td>
<td>visit</td>
</tr>
<tr>
<td>get</td>
<td>meet</td>
<td>wait</td>
</tr>
<tr>
<td>give</td>
<td>met</td>
<td>walk</td>
</tr>
<tr>
<td>go</td>
<td>miss</td>
<td>want</td>
</tr>
<tr>
<td>got</td>
<td>need</td>
<td>warn</td>
</tr>
<tr>
<td>grow</td>
<td>notice</td>
<td>watch</td>
</tr>
<tr>
<td>had</td>
<td>order</td>
<td>welcome</td>
</tr>
<tr>
<td>happen</td>
<td>pass</td>
<td>went</td>
</tr>
<tr>
<td>has</td>
<td>plan</td>
<td>wish</td>
</tr>
<tr>
<td>hate</td>
<td>play</td>
<td>won</td>
</tr>
<tr>
<td>have</td>
<td>point</td>
<td>work</td>
</tr>
<tr>
<td></td>
<td></td>
<td>worry</td>
</tr>
</tbody>
</table>
### Table 8—Grouping 2: List of Verb Tokens Starting with ‘are’, ‘aren’t’, or ‘are not’

<table>
<thead>
<tr>
<th>are</th>
<th>are frighten</th>
<th>are living</th>
</tr>
</thead>
<tbody>
<tr>
<td>are affect</td>
<td>are going</td>
<td>are not</td>
</tr>
<tr>
<td>are clapping</td>
<td>are grow</td>
<td>are not scare</td>
</tr>
<tr>
<td>are decreasing</td>
<td>are growing</td>
<td>are playing</td>
</tr>
<tr>
<td>are die</td>
<td>are hit</td>
<td>are remain</td>
</tr>
<tr>
<td>are feel</td>
<td>are increase</td>
<td>are stay</td>
</tr>
<tr>
<td>are flowing</td>
<td>are laughing</td>
<td>are thank</td>
</tr>
</tbody>
</table>

### Table 9—Grouping 3: List of Verb Tokens Starting with ‘can’, ‘can’t’, or ‘cannot’

<table>
<thead>
<tr>
<th>can ate</th>
<th>can scare</th>
<th>can’t scare</th>
</tr>
</thead>
<tbody>
<tr>
<td>can be</td>
<td>can see</td>
<td>can’t sell</td>
</tr>
<tr>
<td>can became</td>
<td>can shock</td>
<td>can’t shoot</td>
</tr>
<tr>
<td>can become</td>
<td>can shoot</td>
<td>can’t sing</td>
</tr>
<tr>
<td>can brought</td>
<td>can solve</td>
<td>can’t stay</td>
</tr>
<tr>
<td>can came</td>
<td>can stay</td>
<td>can’t take</td>
</tr>
<tr>
<td>can come</td>
<td>can stop</td>
<td>can’t went</td>
</tr>
<tr>
<td>can cross</td>
<td>can swim</td>
<td>cannot accept</td>
</tr>
<tr>
<td>can destroy</td>
<td>can take</td>
<td>cannot afford</td>
</tr>
<tr>
<td>can die</td>
<td>can tell</td>
<td>cannot be</td>
</tr>
<tr>
<td>can do</td>
<td>can turn</td>
<td>cannot be eat</td>
</tr>
<tr>
<td>can drove</td>
<td>can use</td>
<td>cannot buy</td>
</tr>
<tr>
<td>can eat</td>
<td>can visit</td>
<td>cannot came</td>
</tr>
<tr>
<td>can farm</td>
<td>can warm</td>
<td>cannot die</td>
</tr>
<tr>
<td>can feel</td>
<td>can went</td>
<td>cannot do</td>
</tr>
<tr>
<td>can felt</td>
<td>can’t go</td>
<td>cannot drink</td>
</tr>
<tr>
<td>can find</td>
<td>can’t let</td>
<td>cannot earn</td>
</tr>
<tr>
<td>can fly</td>
<td>can’t see</td>
<td>cannot eat</td>
</tr>
<tr>
<td>can frighten</td>
<td>can’t went</td>
<td>cannot enjoy</td>
</tr>
<tr>
<td>can gave</td>
<td>can’t answer</td>
<td>cannot farm</td>
</tr>
<tr>
<td>can get</td>
<td>can’t be</td>
<td>cannot fight</td>
</tr>
<tr>
<td>can give</td>
<td>can’t came</td>
<td>cannot finish</td>
</tr>
<tr>
<td>can go</td>
<td>can’t come</td>
<td>cannot get</td>
</tr>
<tr>
<td>can grow</td>
<td>can’t die</td>
<td>cannot go</td>
</tr>
<tr>
<td>can have</td>
<td>can’t do</td>
<td>cannot grow</td>
</tr>
<tr>
<td>can help</td>
<td>can’t drink</td>
<td>cannot harvest</td>
</tr>
<tr>
<td>can jump</td>
<td>can’t eat</td>
<td>cannot have</td>
</tr>
<tr>
<td>can keep</td>
<td>can’t enjoy</td>
<td>cannot hit</td>
</tr>
<tr>
<td>can kill</td>
<td>can’t farm</td>
<td>cannot hurt</td>
</tr>
<tr>
<td>can leave</td>
<td>can’t get</td>
<td>cannot kill</td>
</tr>
<tr>
<td>can let</td>
<td>can’t go</td>
<td>cannot live</td>
</tr>
<tr>
<td>can live</td>
<td>can’t grow</td>
<td>cannot make...run</td>
</tr>
<tr>
<td>can make</td>
<td>can’t have</td>
<td>cannot plant</td>
</tr>
<tr>
<td>can make...feel</td>
<td>can’t help</td>
<td>cannot return</td>
</tr>
<tr>
<td>can make...run</td>
<td>can’t hit</td>
<td>cannot shoot</td>
</tr>
<tr>
<td>can move</td>
<td>can’t kill</td>
<td>cannot stand</td>
</tr>
<tr>
<td>can plant</td>
<td>can’t live</td>
<td>cannot stay</td>
</tr>
<tr>
<td>can produce</td>
<td>can’t make</td>
<td>cannot stop</td>
</tr>
<tr>
<td>can return</td>
<td>can’t plant</td>
<td>cannot take</td>
</tr>
<tr>
<td>can save</td>
<td>can’t play</td>
<td>cannot went</td>
</tr>
<tr>
<td>can saw</td>
<td>can’t run</td>
<td>cannot work</td>
</tr>
</tbody>
</table>

Table 10—Grouping 4: List of Verb Tokens Starting with ‘could’, ‘couldn’t’, or ‘could not’

| could ate | could not ate | could not went |
| could flew | could not breathe | could pass |
| could gave | could not eat | could return |
| could get | could not farm | could went |
| could go | could not get | couldn’t ate |
| could got | could not kill | couldn’t eat |
| could have | could not took | couldn’t went |

Table 11—Grouping 5: List of Verb Tokens Starting with ‘did’, ‘didn’t’, or ‘did not’

<p>| did | did not hurt | didn’t farm |
| did allow | did not kill | didn’t follow |
| did clap | did not like | didn’t give |
| did die | did not listen | didn’t go |
| did do | did not live | didn’t got |
| did go | did not love | didn’t have |
| did happen | did not need | didn’t hear |
| did have | did not scare | didn’t help |
| did help | did not see | didn’t hit |
| did kill | did not shoot | didn’t kill |
| did laugh | did not stay | didn’t know |
| did listen | did not think | didn’t left |
| did make | did not want | didn’t let |
| did make...go | did not went | didn’t like |
| did need | did return | didn’t listen |
| did not | did say | didn’t listening |
| did not become | did shoot | didn’t love |
| did not bring | did turn | didn’t make |
| did not die | did use | didn’t move |
| did not do | did want | didn’t need |
| did not eat | didn’t | didn’t play |
| did not farm | didn’t fight | didn’t scare |
| did not feel | didn’t go | didn’t see |
| did not finish | didn’t went | didn’t shoot |
| did not get | didn’t allow | didn’t sing |
| did not go | didn’t bear | didn’t stand |</p>
<table>
<thead>
<tr>
<th>did not got</th>
<th>didn’t believe</th>
<th>didn’t take</th>
</tr>
</thead>
<tbody>
<tr>
<td>did not grow</td>
<td>didn’t care</td>
<td>didn’t think</td>
</tr>
<tr>
<td>did not have</td>
<td>didn’t die</td>
<td>didn’t thought</td>
</tr>
<tr>
<td>did not hear</td>
<td>didn’t do</td>
<td>didn’t try</td>
</tr>
<tr>
<td>did not hit</td>
<td>didn’t eat</td>
<td>didn’t use</td>
</tr>
<tr>
<td></td>
<td></td>
<td>didn’t want</td>
</tr>
<tr>
<td></td>
<td></td>
<td>didn’t went</td>
</tr>
</tbody>
</table>

Table 12—Grouping 6: List of Verb Tokens Starting with ‘do’, ‘don’t’, or ‘do not’

<table>
<thead>
<tr>
<th>do</th>
<th>do not love</th>
<th>don’t get</th>
</tr>
</thead>
<tbody>
<tr>
<td>do hope</td>
<td>do not need</td>
<td>don’t give</td>
</tr>
<tr>
<td>do love</td>
<td>do not point</td>
<td>don’t go</td>
</tr>
<tr>
<td>do need</td>
<td>do not seem</td>
<td>don’t have</td>
</tr>
<tr>
<td>do not</td>
<td>do not shoot</td>
<td>don’t kill</td>
</tr>
<tr>
<td>do not agree</td>
<td>do not stand</td>
<td>don’t know</td>
</tr>
<tr>
<td>do not be</td>
<td>do not take</td>
<td>don’t let</td>
</tr>
<tr>
<td>do not come</td>
<td>do not use</td>
<td>don’t like</td>
</tr>
<tr>
<td>do not die</td>
<td>do not want</td>
<td>don’t listen</td>
</tr>
<tr>
<td>do not do</td>
<td>do return</td>
<td>don’t love</td>
</tr>
<tr>
<td>do not feel</td>
<td>do scare</td>
<td>don’t make</td>
</tr>
<tr>
<td>do not felt</td>
<td>do solve</td>
<td>don’t need</td>
</tr>
<tr>
<td>do not give</td>
<td>do welcome</td>
<td>don’t pay</td>
</tr>
<tr>
<td>do not go</td>
<td>don’t</td>
<td>don’t see</td>
</tr>
<tr>
<td>do not have</td>
<td>don’t be</td>
<td>don’t shoot</td>
</tr>
<tr>
<td>do not hit</td>
<td>don’t come</td>
<td>don’t sing</td>
</tr>
<tr>
<td>do not hurt</td>
<td>don’t do</td>
<td>don’t stay</td>
</tr>
<tr>
<td>do not kill</td>
<td>don’t eat</td>
<td>don’t take</td>
</tr>
<tr>
<td>do not listen</td>
<td>don’t feel</td>
<td>don’t use</td>
</tr>
<tr>
<td>do not live</td>
<td>don’t fight</td>
<td>don’t want</td>
</tr>
<tr>
<td></td>
<td></td>
<td>don’t work</td>
</tr>
<tr>
<td></td>
<td></td>
<td>don’t worry</td>
</tr>
</tbody>
</table>

Table 13—Grouping 7: List of Verb Tokens Starting with ‘had’, ‘hadn’t’, or ‘had not’

<table>
<thead>
<tr>
<th>had</th>
<th>had hit</th>
<th>had shot</th>
</tr>
</thead>
<tbody>
<tr>
<td>had done</td>
<td>had kill</td>
<td>hadn’t</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hadn’t go</td>
</tr>
<tr>
<td></td>
<td></td>
<td>hadn’t look</td>
</tr>
</tbody>
</table>

Table 14—Grouping 8: List of Verb Tokens Starting with ‘has’, ‘hasn’t’, or ‘has not’

<table>
<thead>
<tr>
<th>has</th>
<th>has lost</th>
<th>has told</th>
</tr>
</thead>
<tbody>
<tr>
<td>has agree</td>
<td>has not been</td>
<td>has want</td>
</tr>
<tr>
<td>has feel</td>
<td>has shoot</td>
<td>hasn’t been</td>
</tr>
<tr>
<td>has kill</td>
<td>has show</td>
<td>hasn’t kill</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>------------</td>
</tr>
<tr>
<td>has kill</td>
<td>has show</td>
<td>hasn’t kill</td>
</tr>
</tbody>
</table>

**Table 15—Grouping 9: List of Verb Tokens Starting with ‘is’, ‘isn’t’, or ‘is not’**

<table>
<thead>
<tr>
<th>is agree</th>
<th>is gone</th>
<th>is planning</th>
</tr>
</thead>
<tbody>
<tr>
<td>is ask</td>
<td>is grow</td>
<td>is prepare</td>
</tr>
<tr>
<td>is asking</td>
<td>is helping</td>
<td>is require</td>
</tr>
<tr>
<td>is astonish</td>
<td>is holding</td>
<td>is returning</td>
</tr>
<tr>
<td>is believe</td>
<td>is laughing</td>
<td>is scare</td>
</tr>
<tr>
<td>is coming</td>
<td>is like</td>
<td>is shoot</td>
</tr>
<tr>
<td>is die</td>
<td>is living</td>
<td>is showing</td>
</tr>
<tr>
<td>is face</td>
<td>is love</td>
<td>is telling</td>
</tr>
<tr>
<td>is fall</td>
<td>is loving</td>
<td>is thinking</td>
</tr>
<tr>
<td>is feel</td>
<td>is make</td>
<td>is turn</td>
</tr>
<tr>
<td>is flying</td>
<td>is missing</td>
<td>is use</td>
</tr>
<tr>
<td>is frighten</td>
<td>is not go</td>
<td>is using</td>
</tr>
<tr>
<td>is getting</td>
<td>is not going</td>
<td>is want</td>
</tr>
<tr>
<td>is give</td>
<td>is not love</td>
<td>is welcome</td>
</tr>
<tr>
<td>is go</td>
<td>is not scare</td>
<td>is working</td>
</tr>
<tr>
<td>is going</td>
<td>is not working</td>
<td>is worry</td>
</tr>
<tr>
<td></td>
<td></td>
<td>isn’t</td>
</tr>
</tbody>
</table>

**Table 16—Grouping 10: List of Verb Tokens Starting with ‘make’**

<table>
<thead>
<tr>
<th>make...cannot plant</th>
<th>make...go</th>
</tr>
</thead>
<tbody>
<tr>
<td>make...absorb</td>
<td>make...did not feel</td>
</tr>
<tr>
<td>make...appear</td>
<td>make...did not have</td>
</tr>
<tr>
<td>make...are</td>
<td>make...did not have</td>
</tr>
<tr>
<td>make...became</td>
<td>make...did not plant</td>
</tr>
<tr>
<td>make...become</td>
<td>make...did suffer</td>
</tr>
<tr>
<td>make...can grew</td>
<td>make...didn't have</td>
</tr>
<tr>
<td>make...cannot farm</td>
<td>make...die</td>
</tr>
<tr>
<td>make...can plant</td>
<td>make...farm</td>
</tr>
<tr>
<td>make...can't farm</td>
<td>make...feel</td>
</tr>
<tr>
<td>make...cannot farm</td>
<td>make...feel, felt</td>
</tr>
<tr>
<td>make...cannot grow</td>
<td>make...felt</td>
</tr>
<tr>
<td>make...cannot live</td>
<td>make...get</td>
</tr>
</tbody>
</table>

**Table 17—Grouping 11: List of Verb Tokens Starting with ‘may’ or ‘may not’**

<table>
<thead>
<tr>
<th>may be</th>
<th>may have</th>
<th>may lost</th>
</tr>
</thead>
<tbody>
<tr>
<td>may escape</td>
<td>may help</td>
<td>may need</td>
</tr>
<tr>
<td>may go</td>
<td>may let</td>
<td>may not</td>
</tr>
</tbody>
</table>
Table 18—Grouping 12: List of Verb Tokens Starting with ‘must’, ‘mustn’t’, or ‘must not’

<table>
<thead>
<tr>
<th>must be go</th>
<th>must go</th>
<th>must pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>must become</td>
<td>must got</td>
<td>must saw</td>
</tr>
<tr>
<td>must cross</td>
<td>must have</td>
<td>must share</td>
</tr>
<tr>
<td>must do</td>
<td>must lead</td>
<td>must take</td>
</tr>
<tr>
<td>must eat</td>
<td>must leave</td>
<td>must went</td>
</tr>
<tr>
<td>must get</td>
<td>must not eat</td>
<td>mustn’t eat</td>
</tr>
</tbody>
</table>

Table 19—Grouping 13: List of Verb Tokens Starting with ‘see’

<table>
<thead>
<tr>
<th>see</th>
<th>see...happen</th>
<th>see...make</th>
</tr>
</thead>
<tbody>
<tr>
<td>see...fly</td>
<td>see...is</td>
<td>see...were</td>
</tr>
</tbody>
</table>

Table 20—Grouping 14: List of Verb Tokens Starting with ‘should’, ‘shouldn’t’, or ‘should not’

<table>
<thead>
<tr>
<th>should did</th>
<th>should let</th>
<th>should pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>should drove</td>
<td>should met</td>
<td>should shot</td>
</tr>
<tr>
<td>should gave</td>
<td>should not ate</td>
<td>should took</td>
</tr>
<tr>
<td>should give</td>
<td>should not kill</td>
<td>should went</td>
</tr>
<tr>
<td>should keep</td>
<td>should not shot</td>
<td>shouldn’t did</td>
</tr>
<tr>
<td></td>
<td></td>
<td>shouldn’t shot</td>
</tr>
</tbody>
</table>

Table 21—Grouping 15: List of Verb Tokens Starting with ‘will’, ‘won’t’, or ‘will not’

<table>
<thead>
<tr>
<th>will be</th>
<th>will lead</th>
<th>will not stay</th>
</tr>
</thead>
<tbody>
<tr>
<td>will become</td>
<td>will let</td>
<td>will not use</td>
</tr>
<tr>
<td>will bring</td>
<td>will make</td>
<td>will punish</td>
</tr>
<tr>
<td>will call</td>
<td>will make...grow</td>
<td>will return</td>
</tr>
<tr>
<td>will come</td>
<td>will meet</td>
<td>will shot</td>
</tr>
<tr>
<td>will die</td>
<td>will not be</td>
<td>will suffer</td>
</tr>
<tr>
<td>will do</td>
<td>will not come</td>
<td>will take</td>
</tr>
<tr>
<td>will escape</td>
<td>will not die</td>
<td>will think</td>
</tr>
<tr>
<td>will face</td>
<td>will not do</td>
<td>will told</td>
</tr>
<tr>
<td>will feel</td>
<td>will not drop</td>
<td>will turn</td>
</tr>
<tr>
<td>will fly</td>
<td>will not get</td>
<td>will use</td>
</tr>
<tr>
<td>will get</td>
<td>will not go</td>
<td>will worry</td>
</tr>
<tr>
<td>will give</td>
<td>will not have</td>
<td>won’t be</td>
</tr>
<tr>
<td>will go</td>
<td>will not hit</td>
<td>won’t die</td>
</tr>
<tr>
<td>Verb Token 1</td>
<td>Verb Token 2</td>
<td>Verb Token 3</td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>will got</td>
<td>will not hurt</td>
<td>won't have</td>
</tr>
<tr>
<td>will grow</td>
<td>will not kill</td>
<td>won't hurt</td>
</tr>
<tr>
<td>will happen</td>
<td>will not let...die</td>
<td>won't kill</td>
</tr>
<tr>
<td>will have</td>
<td>will not make</td>
<td>won't let</td>
</tr>
<tr>
<td>will help</td>
<td>will not return</td>
<td>won't like</td>
</tr>
<tr>
<td>will kick</td>
<td>will not shoot</td>
<td>won't scare</td>
</tr>
<tr>
<td>will kill</td>
<td>will not shoot</td>
<td>won't shoot</td>
</tr>
</tbody>
</table>

Table 22—Grouping 16: List of Verb Tokens Starting with ‘would’, ‘wouldn’t’, or ‘would not’

<table>
<thead>
<tr>
<th>Verb Token 1</th>
<th>Verb Token 2</th>
<th>Verb Token 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>would be</td>
<td>would go</td>
<td>would not felt</td>
</tr>
<tr>
<td>would became</td>
<td>would got</td>
<td>would not kill</td>
</tr>
<tr>
<td>would brought</td>
<td>would have</td>
<td>would not shot</td>
</tr>
<tr>
<td>would die</td>
<td>would kill</td>
<td>would went</td>
</tr>
<tr>
<td>would do</td>
<td>would let</td>
<td>wouldn’t hurt</td>
</tr>
<tr>
<td>would get</td>
<td>would make</td>
<td>wouldn’t kill</td>
</tr>
<tr>
<td></td>
<td></td>
<td>wouldn’t shoot</td>
</tr>
</tbody>
</table>

Table 23—Grouping 17: List of Verb Tokens Starting with ‘were’, ‘weren’t’, or ‘were not’

<table>
<thead>
<tr>
<th>Verb Token 1</th>
<th>Verb Token 2</th>
<th>Verb Token 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>were</td>
<td>were laughing</td>
<td>were require</td>
</tr>
<tr>
<td>were arrive</td>
<td>were left</td>
<td>were say</td>
</tr>
<tr>
<td>were die</td>
<td>were living</td>
<td>were shoot</td>
</tr>
<tr>
<td>were felt</td>
<td>were not</td>
<td>were shot</td>
</tr>
<tr>
<td>were frighten</td>
<td>were not go</td>
<td>were smiling</td>
</tr>
<tr>
<td>were gone</td>
<td>were not listen</td>
<td>were suffer</td>
</tr>
<tr>
<td>were grow</td>
<td>were not listening</td>
<td>were thank</td>
</tr>
<tr>
<td>were help</td>
<td>were not scare</td>
<td>were try</td>
</tr>
<tr>
<td>were laugh</td>
<td>were playing</td>
<td>were turn</td>
</tr>
<tr>
<td></td>
<td></td>
<td>were walking</td>
</tr>
<tr>
<td></td>
<td></td>
<td>were worry</td>
</tr>
</tbody>
</table>

Table 24—Grouping 18: List of Verb Tokens Starting with ‘does’, ‘doesn’t’, or ‘does not’

<table>
<thead>
<tr>
<th>Verb Token 1</th>
<th>Verb Token 2</th>
<th>Verb Token 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>does help</td>
<td>does not need</td>
<td>doesn’t have</td>
</tr>
<tr>
<td>does love</td>
<td>does not work</td>
<td>doesn’t kill</td>
</tr>
<tr>
<td>does not</td>
<td>does return</td>
<td>doesn’t know</td>
</tr>
<tr>
<td>does not go</td>
<td>does show</td>
<td>doesn’t like</td>
</tr>
<tr>
<td>does not hurt</td>
<td>doesn’t die</td>
<td>doesn’t love</td>
</tr>
<tr>
<td>does not kill</td>
<td>doesn’t eat</td>
<td>doesn’t think</td>
</tr>
<tr>
<td>does not live</td>
<td>doesn’t go</td>
<td>doesn’t want</td>
</tr>
</tbody>
</table>
Table 25—Grouping 19: List of Verb Tokens Starting with ‘have’, ‘haven’t’, or ‘have not’

<table>
<thead>
<tr>
<th>Verb</th>
<th>Verb</th>
<th>Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>have</td>
<td>have got</td>
<td>have say</td>
</tr>
<tr>
<td>have affect</td>
<td>have grow</td>
<td>have shoot</td>
</tr>
<tr>
<td>have been</td>
<td>have hit</td>
<td>have think</td>
</tr>
<tr>
<td>have been dry</td>
<td>have kill</td>
<td>have thought</td>
</tr>
<tr>
<td>have clear</td>
<td>have left</td>
<td>have want</td>
</tr>
<tr>
<td>have done</td>
<td>have make</td>
<td>haven’t get</td>
</tr>
<tr>
<td>have give</td>
<td>have not</td>
<td>haven’t got</td>
</tr>
<tr>
<td>have given</td>
<td>have not kill</td>
<td>haven’t listen</td>
</tr>
<tr>
<td>have gone</td>
<td>have not listen</td>
<td>haven’t shoot</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Verb</th>
<th>Verb</th>
<th>Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>haven’t</td>
<td>tell</td>
<td></td>
</tr>
</tbody>
</table>

Table 26—Grouping 20: List of Verb Tokens Starting with ‘was’, ‘wasn’t’, or ‘was not’

<table>
<thead>
<tr>
<th>Verb</th>
<th>Verb</th>
<th>Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>was</td>
<td>was help</td>
<td>was returning</td>
</tr>
<tr>
<td>was accept</td>
<td>was hope</td>
<td>was say</td>
</tr>
<tr>
<td>was agree</td>
<td>was kept</td>
<td>was scare</td>
</tr>
<tr>
<td>was alive</td>
<td>was kill</td>
<td>was see</td>
</tr>
<tr>
<td>was annoy</td>
<td>was laugh</td>
<td>was shining</td>
</tr>
<tr>
<td>was arrive</td>
<td>was laughing</td>
<td>was shoot</td>
</tr>
<tr>
<td>was ask</td>
<td>was left</td>
<td>was shooting</td>
</tr>
<tr>
<td>was ate</td>
<td>was listen</td>
<td>was shot</td>
</tr>
<tr>
<td>was beg</td>
<td>was live</td>
<td>was shout</td>
</tr>
<tr>
<td>was broken</td>
<td>was look</td>
<td>was stay</td>
</tr>
<tr>
<td>was climb</td>
<td>was looking</td>
<td>was surprise</td>
</tr>
<tr>
<td>was come</td>
<td>was love</td>
<td>was telling</td>
</tr>
<tr>
<td>was damage</td>
<td>was lying</td>
<td>was thank</td>
</tr>
<tr>
<td>was daydreaming</td>
<td>was make</td>
<td>was think</td>
</tr>
<tr>
<td>was decide</td>
<td>was miss</td>
<td>was thinking</td>
</tr>
<tr>
<td>was destroy</td>
<td>was need</td>
<td>was told</td>
</tr>
<tr>
<td>was die</td>
<td>was not</td>
<td>was travel</td>
</tr>
<tr>
<td>was disturb</td>
<td>was not die</td>
<td>was turn</td>
</tr>
<tr>
<td>was do</td>
<td>was not go</td>
<td>was use</td>
</tr>
<tr>
<td>was eating</td>
<td>was not hit</td>
<td>was want</td>
</tr>
<tr>
<td>was embarrass</td>
<td>was not hurt</td>
<td>was went</td>
</tr>
<tr>
<td>was fly</td>
<td>was not kill</td>
<td>was worry</td>
</tr>
<tr>
<td>was flying</td>
<td>was not want</td>
<td>wasn’t</td>
</tr>
<tr>
<td>was getting</td>
<td>was planning</td>
<td>wasn’t buy</td>
</tr>
<tr>
<td>was go</td>
<td>was promise</td>
<td>wasn’t die</td>
</tr>
<tr>
<td>was going</td>
<td>was repeat</td>
<td>wasn’t hurt</td>
</tr>
<tr>
<td>was gone</td>
<td>was require</td>
<td>wasn’t kill</td>
</tr>
</tbody>
</table>
In addition to displaying students’ use of conventional verb tokens, for example “became”, “buy”, and “may think”, tables 7-27 also demonstrate their inaccurate use of verbs. As shown above, students used auxiliary verbs to signal their use of past tense, for example “were grow”, and their use of non-past tense, for example “is go”. This may exemplify learners’ use of verb-raising to mark tense, as discussed in section 4.6.1.2 of chapter 4. Moreover, students’ interlanguage or developing use of past tense, such as “didn’t went”, has been revealed. Some regular past verb tokens in grouping 1, verb tokens starting with “had”, “hadn’t”, or “had not” in grouping 7, “has”, “hasn’t”, or “has not” in grouping 8, “make” in grouping 10, and “have”, “haven’t”, or “have not” in grouping 19 were transcribed based on the raters’ perception of what the speakers uttered at that moment. However, as explained in chapter 4, students’ spoken use of regular past tokens and irregular past tokens with no vowel change from their base form could not be ensured, because of their ambiguous phonetic representation in speech. Therefore, these tokens were excluded from the study’s past tense usage analysis. Verb tokens starting with “make” in grouping 10, “see” in grouping 13, and “saw” in grouping 21 have been particularly presented for reference, because they have accompanying bare infinitive verbs following which do not require any obligatory tense marking.

Table 27—Grouping 21: List of Verb Tokens Starting with ‘saw’

<table>
<thead>
<tr>
<th>saw</th>
<th>saw...go</th>
<th>saw...say</th>
</tr>
</thead>
</table>

was grow | was return | wasn’t like |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>wasn’t love</td>
</tr>
<tr>
<td></td>
<td></td>
<td>wasn’t scare</td>
</tr>
</tbody>
</table>
5.8 ANOVA on Different Types of Verbs Included and Excluded in the Data Analysis

As specified in chapter 4, not all verb tokens used by the participants were analysed. Only some past tense verb tokens were included in the analysis and others were excluded, due to the issue of result validity. The ANOVA analyses presented in this section describe how the participants’ use of the verb tokens, included as well as excluded in the data analysis, differed across the different feedback groups and test time points. The verb tokens were sorted into 9 different verb types: 1) developing past form, or interlanguage, e.g. “didn’t took”, 2) irregular past form, e.g. “went”, 3) non-past form, e.g. “buy”, 4) modal verbs, e.g. “may give”, 5) same past form, or verbs with the same past and base form, e.g. “put”, 6) verbs in monologue or dialogue, with no obligatory use of past tense, 7) verbless L1 transfer, e.g. “He very angry”, 8) no vowel change past form, or past form with no vowel change from its base form, e.g. “make”, and 9) regular past form, e.g. “walk”. As specified in chapter 4, only 1) and 2) were included in the study’s past tense usage analysis.

Tables 28-36 show the main effects and their interaction effects of the aforementioned 9 variables across time and feedback groups. Same as tables 3 and 4, the two main effects in the analyses of the 9 variables were the effects of the different feedback treatments (corrective recasts, normal recasts, and content-only control feedback) and the different time points (pretest, feedback session, immediate posttest, and delayed posttest). The time effect measured in the two-way mixed ANOVA analyses on the 9 variables considered the feedback session time point, making four time points in total. This deviates from the three time points (pretest, immediate posttest, and delayed posttest) employed in the two-way mixed and one-way ANOVA analyses on students’ past tense usage ratio presented in tables 3 and 4 respectively. The exclusion of the feedback session time point in tables 3 and 4 is due to the fact
that students’ past tense usage and the effectiveness of the different feedbacks they received across time were measured based on their pre and post-treatment performance. The feedback session time point in tables 28-36 is included because the 9 variables represent students’ use of a wide range of form in their narratives—past tense usage, non-past tense usage, and ambiguous form. These 9 variables only provide an insight into what students actually used and how they differed across different groups and time, but do not assess students’ improving or deteriorating use of these forms.

The interpretation of the main effects is the difference in means when levels (3 feedback groups) of one factor (group) collapse over levels (4 time points) of the other factor (time). A simple main effect is the main effect of one factor at a given level of a second factor; for example, at the pretest time level, whether or not there is a group effect among the three levels of the group factor. Accordingly, the interaction effect between the group and time effects is a change in the simple main effect of one variable over levels of the second variable; for example, whether or not there is a difference between the differences among the three groups at the different levels of time. The main and interaction effects convey three different meanings. The main effect results provide an overview of the general effect of the factor by ignoring or collapsing levels of other factors. The interaction effect results channel how the simple effect of one factor changes at different levels of another factor. It is therefore possible for a set of data having only significant interaction effect but no main effect, or just one main effect significant.

The ANOVA results overall deliver if the main effects and interaction effect are statistically significant after statistical hypothesis testing. For example, in the measure of the first variable “developing past form”, the insignificant group effect
showed in general there was no statistical difference among the three groups when all levels of time were collapsed. The significant time effect on the other hand demonstrated in general there were differences among different time points when the three groups were collapsed. The significant interaction effect indicated that the differences of the differences among the three groups at each time point were significantly different.

The following 9 tables investigate the time effect (within-subjects), group effect (between subjects) and interaction effect on the 9 aforementioned variables. There are all together 4 levels (pretest, feedback session, immediate posttest, delayed posttest) for the time effect, and 3 levels (corrective recasts, normal recasts, content-only control feedback) for the group effect. The significance level is predetermined at 0.05 and 0.01. Tables 28-36 are the major references, and the accompanying figures are drawn only to aid understanding.
Table 29– Results of Two-way Mixed ANOVA on the Outcome Measure ‘Developing Past Form’

<table>
<thead>
<tr>
<th>Group</th>
<th>Total (n=89)</th>
<th>Control (C) (n = 31)</th>
<th>Normal recasts (NR) (n = 29)</th>
<th>Corrective recasts (CR) (n = 29)</th>
<th>Group effect (Greenhouse-Geisser)</th>
<th>Time effect (Greenhouse-Geisser)</th>
<th>Interaction effect (Group*Time) (Greenhouse-Geisser)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>F(2,86) = 1.180 p = 0.312 $\eta^2 = 0.027$</td>
<td>F (2.518,216.564) = 12.434 p &lt; 0.001** $\eta^2 = 0.126$</td>
<td>F (5.036,216.564) = 3.369 p = 0.006** $\eta^2 = 0.073$</td>
</tr>
<tr>
<td>Pretest</td>
<td>0.71 1.1</td>
<td>0.90 1.1</td>
<td>0.66 1.3</td>
<td>0.55 0.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback</td>
<td>0.34 0.6</td>
<td>0.32 0.7</td>
<td>0.38 0.6</td>
<td>0.31 0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate Posttest</td>
<td>1.24 1.5</td>
<td>0.87 1.1</td>
<td>1.17 1.4</td>
<td>1.69 1.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed Posttest</td>
<td>1.16 1.7</td>
<td>0.65 1.3</td>
<td>1.79 2.1</td>
<td>1.07 1.3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < 0.01; $\eta^2$ denotes the estimate of partial eta square as a measure of effect size; Greenhouse-Geisser correction was applied to produce a valid F-ratio.
Figure 2—Graphical Results of Two-way Mixed ANOVA on the Outcome Measure ‘Developing Past Form’

Note: Time 1, 2, 3 and 4 represents pre-test, feedback, immediate posttest and delayed posttest respectively.
Table 30– Results of Two-way Mixed ANOVA on the Outcome Measure ‘Irregular Past Form’

<table>
<thead>
<tr>
<th>Group</th>
<th>Total (n=89)</th>
<th>Control (C) (n=31)</th>
<th>Normal recasts (NR) (n = 29)</th>
<th>Corrective recasts (CR) (n = 29)</th>
<th>Group effect (Greenhouse-Geisser)</th>
<th>Time effect</th>
<th>Interaction effect (Group*Time) (Greenhouse-Geisser)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>F(2,86) = 5.450 p = 0.006** η² = 0.112</td>
<td>F (2.766,237.870) = 30.738 p &lt; 0.001** η² = 0.263</td>
<td>F (5.532,237.870) = 3.309 p = 0.005** η² = 0.071</td>
</tr>
<tr>
<td>Pretest</td>
<td>6.22 4.6</td>
<td>6.81 4.0</td>
<td>5.21 5.0</td>
<td>6.62 4.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Feedback</td>
<td>6.82 5.1</td>
<td>5.03 4.6</td>
<td>7.14 5.7</td>
<td>8.41 4.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Immediate Posttest</td>
<td>11.78 6.4</td>
<td>8.58 5.6</td>
<td>12.38 5.8</td>
<td>14.59 6.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed Posttest</td>
<td>9.17 6.2</td>
<td>6.94 5.1</td>
<td>9.72 5.8</td>
<td>11.00 7.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < 0.01; η² denotes the estimate of partial eta square as a measure of effect size; Greenhouse-Geisser correction was applied to produce a valid F-ratio.
Figure 3—Graphical Results of Two-way Mixed ANOVA on the Outcome Measure ‘Irregular Past Form’

Note: Time 1, 2, 3 and 4 represents pre-test, feedback, immediate posttest and delayed posttest respectively.
Table 31 - Results of Two-way Mixed ANOVA on the Outcome Measure 'Non-Past Form'

<table>
<thead>
<tr>
<th>Group</th>
<th>Total (n=89)</th>
<th>Control (C) (n = 31)</th>
<th>Normal recasts (NR) (n = 29)</th>
<th>Corrective recasts (CR) (n = 29)</th>
<th>Group effect (Greenhouse-Geisser)</th>
<th>Time effect (Greenhouse-Geisser)</th>
<th>Interaction effect (Group*Time) (Greenhouse-Geisser)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Pretest</td>
<td>9.78</td>
<td>5.4</td>
<td>8.81</td>
<td>5.0</td>
<td>9.86</td>
<td>6.3</td>
<td>10.72</td>
</tr>
<tr>
<td>Feedback</td>
<td>3.08</td>
<td>5.2</td>
<td>8.84</td>
<td>5.2</td>
<td>0.00</td>
<td>0.0</td>
<td>0.00</td>
</tr>
<tr>
<td>Immediate Posttest</td>
<td>6.18</td>
<td>4.9</td>
<td>8.65</td>
<td>5.4</td>
<td>5.07</td>
<td>5.0</td>
<td>4.66</td>
</tr>
<tr>
<td>Delayed Posttest</td>
<td>6.12</td>
<td>4.6</td>
<td>7.81</td>
<td>4.4</td>
<td>4.90</td>
<td>4.3</td>
<td>5.55</td>
</tr>
</tbody>
</table>

p < 0.05; ** p < 0.01; η² denotes the estimate of partial eta square as a measure of effect size; Greenhouse-Geisser correction was applied to produce a valid F-ratio.
Figure 4—Graphical Results of Two-way Mixed ANOVA on the Outcome Measure ‘Non-Past Form’

Note: Time 1, 2, 3 and 4 represents pre-test, feedback, immediate posttest and delayed posttest respectively.
Table 32– Results of Two-way Mixed ANOVA on the Outcome Measure ‘Modal Verbs’

<table>
<thead>
<tr>
<th>Group</th>
<th>Total (n=89)</th>
<th>Control (C) (n = 31)</th>
<th>Normal recasts (NR) (n = 29)</th>
<th>Corrective recasts (CR) (n = 29)</th>
<th>Group effect</th>
<th>Time effect</th>
<th>Interaction effect (Group*Time)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>Mean SD</td>
<td>F(2,86) = 1.882</td>
<td>F(3,258) = 5.293</td>
<td>F (6, 258) = 5.790</td>
</tr>
<tr>
<td>Pretest</td>
<td>3.15 2.8</td>
<td>2.61 2.1</td>
<td>3.34 2.9</td>
<td>3.52 3.3</td>
<td>p = 0.159</td>
<td>p = 0.001**</td>
<td>p &lt; 0.001**</td>
</tr>
<tr>
<td>Feedback</td>
<td>1.89 2.4</td>
<td>3.65 3.0</td>
<td>0.83 1.2</td>
<td>1.07 1.4</td>
<td>r^2 = 0.042</td>
<td>r^2 = 0.058</td>
<td>r^2 = 0.119</td>
</tr>
<tr>
<td>Immediate Posttest</td>
<td>2.63 2.5</td>
<td>3.29 2.9</td>
<td>2.83 2.7</td>
<td>1.72 1.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Delayed Posttest</td>
<td>2.37 2.3</td>
<td>2.10 1.5</td>
<td>2.41 3.0</td>
<td>2.62 2.2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

p < 0.05; ** p < 0.01; \( \eta^2 \) denotes the estimate of partial eta square as a measure of effect size.
Figure 5—Graphical Results of Two-way Mixed ANOVA on the Outcome Measure ‘Modal Verbs’

Note: Time 1, 2, 3 and 4 represents pre-test, feedback, immediate posttest and delayed posttest respectively.
Table 33—Results of Two-way Mixed ANOVA on the Outcome Measure ‘Same Past Form’

<table>
<thead>
<tr>
<th>Group</th>
<th>Total (n=89)</th>
<th>Control (C) (n = 31)</th>
<th>Normal recasts (NR) (n = 29)</th>
<th>Corrective recasts (CR) (n = 29)</th>
<th>Group effect (Greenhouse-Geisser)</th>
<th>Time effect (Greenhouse-Geisser)</th>
<th>Interaction effect (Group*Time) (Greenhouse-Geisser)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Pretest</td>
<td>0.19</td>
<td>0.5</td>
<td>0.06</td>
<td>0.2</td>
<td>0.28</td>
<td>0.6</td>
<td>0.24</td>
</tr>
<tr>
<td>Feedback</td>
<td>0.21</td>
<td>0.5</td>
<td>0.29</td>
<td>0.6</td>
<td>0.21</td>
<td>0.5</td>
<td>0.14</td>
</tr>
<tr>
<td>Immediate Posttest</td>
<td>0.27</td>
<td>0.6</td>
<td>0.29</td>
<td>0.7</td>
<td>0.28</td>
<td>0.5</td>
<td>0.24</td>
</tr>
<tr>
<td>Delayed Posttest</td>
<td>0.43</td>
<td>0.7</td>
<td>0.48</td>
<td>0.9</td>
<td>0.45</td>
<td>0.7</td>
<td>0.34</td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < 0.01; η² denotes the estimate of partial eta square as a measure of effect size; Greenhouse-Geisser correction was applied to produce a valid F-ratio.
Estimated Marginal Means of MEASURE_1

Figure 6—Graphical Results of Two-way Mixed ANOVA on the Outcome Measure ‘Same Past Form’

Note: Time 1, 2, 3 and 4 represents pre-test, feedback, immediate posttest and delayed posttest respectively.
Table 34 – Results of Two-way Mixed ANOVA on the Outcome Measure ‘Verbs in Monologue or Dialogue’

<table>
<thead>
<tr>
<th>Test</th>
<th>Total (n=89)</th>
<th>Control (C) (n = 31)</th>
<th>Normal recasts (NR) (n = 29)</th>
<th>Corrective recasts (CR) (n = 29)</th>
<th>Group effect (Greenhouse-Geisser)</th>
<th>Time effect (Greenhouse-Geisser)</th>
<th>Interaction effect (Group*Time) (Greenhouse-Geisser)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Pretest</td>
<td>0.71</td>
<td>2.0</td>
<td>1.23</td>
<td>3.1</td>
<td>0.59</td>
<td>1.3</td>
<td>0.28</td>
</tr>
<tr>
<td>Feedback</td>
<td>0.47</td>
<td>1.2</td>
<td>0.81</td>
<td>1.7</td>
<td>0.14</td>
<td>0.4</td>
<td>0.45</td>
</tr>
<tr>
<td>Immediate Posttest</td>
<td>0.69</td>
<td>1.8</td>
<td>1.23</td>
<td>2.6</td>
<td>0.31</td>
<td>1.1</td>
<td>0.48</td>
</tr>
<tr>
<td>Delayed Posttest</td>
<td>0.44</td>
<td>1.4</td>
<td>0.87</td>
<td>2.2</td>
<td>0.21</td>
<td>0.8</td>
<td>0.21</td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < 0.01; η² denotes the estimate of partial eta square as a measure of effect size; Greenhouse-Geisser correction was applied to produce a valid F-ratio.
Figure 7—Graphical Results of Two-way Mixed ANOVA on the Outcome Measure ‘Verbs in Monologue or Dialogue’

Note: Time 1, 2, 3 and 4 represents pre-test, feedback, immediate posttest and delayed posttest respectively.
Table 35– Results of Two-way Mixed ANOVA on the Outcome Measure ‘Verbless L1 Transfer’

<table>
<thead>
<tr>
<th>Test</th>
<th>Group</th>
<th>Total (n=89)</th>
<th>Control (C) (n=31)</th>
<th>Normal recasts (NR) (n=29)</th>
<th>Corrective recasts (CR) (n=29)</th>
<th>Group effect</th>
<th>Time effect</th>
<th>Interaction effect (Group*Time)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pretest</td>
<td>Total (n=89)</td>
<td>0.42 Mean</td>
<td>0.6 SD</td>
<td>0.42 Mean</td>
<td>0.6 SD</td>
<td>0.52 Mean</td>
<td>0.7 SD</td>
<td>0.31 Mean</td>
</tr>
<tr>
<td>Feedback</td>
<td>Control (C) (n=31)</td>
<td>0.26 Mean</td>
<td>0.6 SD</td>
<td>0.68 Mean</td>
<td>0.9 SD</td>
<td>0.03 Mean</td>
<td>0.2 SD</td>
<td>0.03 Mean</td>
</tr>
<tr>
<td>Immediate Posttest</td>
<td>Normal recasts (NR) (n=29)</td>
<td>0.33 Mean</td>
<td>0.45 SD</td>
<td>0.17 Mean</td>
<td>0.5 SD</td>
<td>0.34 Mean</td>
<td>0.7 SD</td>
<td>F(2,86) = 7.949 ( p = 0.001^{**} ) ( \eta^2 = 0.156 )</td>
</tr>
<tr>
<td>Delayed Posttest</td>
<td>Corrective recasts (CR) (n=29)</td>
<td>0.24 Mean</td>
<td>0.55 SD</td>
<td>0.03 Mean</td>
<td>0.2 SD</td>
<td>0.10 Mean</td>
<td>0.3 SD</td>
<td>F (3,258) = 2.015 ( p = 0.112 ) ( \eta^2 = 0.023 )</td>
</tr>
</tbody>
</table>

\* p < 0.05; \** p < 0.01; \( \eta^2 \) denotes the estimate of partial eta square as a measure of effect size.
Figure 8—Graphical Results of Two-way Mixed ANOVA on the Outcome Measure ‘Verbless L1 Transfer’

Note: Time 1, 2, 3 and 4 represents pre-test, feedback, immediate posttest and delayed posttest respectively.
Table 36 – Results of Two-way Mixed ANOVA on the Outcome Measure ‘No Vowel Change Past Form’

<table>
<thead>
<tr>
<th>Group</th>
<th>Total (n=89)</th>
<th>Control (C) (n=31)</th>
<th>Normal recasts (NR) (n=29)</th>
<th>Corrective recasts (CR) (n=29)</th>
<th>Group effect</th>
<th>Time effect</th>
<th>Interaction effect (Group*Time)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Pretest</td>
<td>1.73</td>
<td>1.9</td>
<td>1.32</td>
<td>1.7</td>
<td>1.66</td>
<td>1.7</td>
<td>2.24</td>
</tr>
<tr>
<td>Feedback</td>
<td>2.03</td>
<td>1.8</td>
<td>3.19</td>
<td>1.6</td>
<td>1.34</td>
<td>1.7</td>
<td>1.48</td>
</tr>
<tr>
<td>Immediate Posttest</td>
<td>1.30</td>
<td>1.6</td>
<td>1.84</td>
<td>1.8</td>
<td>1.21</td>
<td>1.6</td>
<td>0.83</td>
</tr>
<tr>
<td>Delayed Posttest</td>
<td>1.12</td>
<td>1.6</td>
<td>1.16</td>
<td>1.5</td>
<td>1.07</td>
<td>1.3</td>
<td>1.14</td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < 0.01; η² denotes the estimate of partial eta square as a measure of effect size.
Figure 9—Graphical Results of Two-way Mixed ANOVA on the Outcome Measure ‘No Vowel Change Past Form’

Note: Time 1, 2, 3 and 4 represents pre-test, feedback, immediate posttest and delayed posttest respectively.
Table 37 – Results of Two-way Mixed ANOVA on the Outcome Measure ‘Regular Past Form’

<table>
<thead>
<tr>
<th>Group</th>
<th>Total (n=89)</th>
<th>Control (C) (n=31)</th>
<th>Normal recasts (NR) (n=29)</th>
<th>Corrective recasts (CR) (n=29)</th>
<th>Group effect</th>
<th>Time effect</th>
<th>Interaction effect (Group*Time)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Test</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>5.56</td>
<td>2.6</td>
<td>5.00</td>
<td>2.4</td>
<td>5.76</td>
<td>2.1</td>
<td>5.97</td>
</tr>
<tr>
<td>Feedback</td>
<td>2.66</td>
<td>2.8</td>
<td>5.42</td>
<td>2.4</td>
<td>0.66</td>
<td>0.9</td>
<td>1.72</td>
</tr>
<tr>
<td>Immediate Posttest</td>
<td>4.16</td>
<td>2.3</td>
<td>4.45</td>
<td>2.2</td>
<td>3.86</td>
<td>2.7</td>
<td>4.14</td>
</tr>
<tr>
<td>Delayed Posttest</td>
<td>4.69</td>
<td>2.7</td>
<td>4.90</td>
<td>3.1</td>
<td>4.86</td>
<td>2.5</td>
<td>4.28</td>
</tr>
</tbody>
</table>

* p < 0.05; ** p < 0.01; η² denotes the estimate of partial eta square as a measure of effect size.
Figure 10—Graphical Results of Two-way Mixed ANOVA on the Outcome Measure ‘Regular Past Form’

Note: Time 1, 2, 3 and 4 represents pre-test, feedback, immediate posttest and delayed posttest respectively.
From tables 28-36, variables with their group and time levels contributed the statistically significant main and interaction effects were irregular past form, non-past form, and regular past form. Variables with only the time main effect and interaction effect significant were developing past form, modal verbs, and no vowel change past form. The verbless L1 transfer variable was the only variable with significant group effect and interaction effect but no time effect. The same past form variable was the only variable with only time effect significant. Variables with no group and time levels contributing any significant main and interaction effects were verbs in monologue or dialogue.

5.9 Feedback Session Variables

In the normal recast and corrective recast feedback sessions, only verbs fulfilling the result validity criteria, as set in chapter 4, and therefore intended to be analysed were targeted by either normal recasts or corrective recasts. This was to attempt an accurate test on the effectiveness of the two aforementioned feedback types on the participants' use of past tense across time. However, the test of the effectiveness of the feedback types may not unveil the different variables occurred in the feedback sessions, which may have influenced the subsequent test results. The following analyses describe how the variables appeared in the normal and corrective recast sessions differed across the different feedback types.

First of all, table 37 shows that students' use of past tense ratio was significantly higher in the normal and corrective recast sessions than that in the content-only feedback control session. This gives extra information that students in the two recast sessions also used more past tense than those in the control session. Basically, all the figures below provide graphical overviews. Major interpretations are
based on the test results in tables; the figures mainly serve as assistance. The figures however unveil the existence of some outliers and extreme values. As shown in figure 11, the control feedback group got the lowest use of past tense ratio, and the normal and corrective recast feedback groups got the similar and higher use of past tense ratio. Moreover, the corrective recast feedback group got the lowest data variance, but the control feedback group got the highest data variance. However, there were outliers and extreme values in the normal recast and corrective recast groups. It seems that some individuals had extraordinary performance. The circles indicate the outliers which had values ranged from median +/- 1.5*IQR to median +/- 3*IQR. IQR represents the inter-quartile range. Case 40 in the normal recast group had identity code NR3.36 in the dataset; Cases 9, 11 and 14 in the corrective recast group had identity codes CR2.37.SR, CR3.45, and CR1.55.SR respectively in the dataset.

Regarding this, the ANOVA test and the post hoc comparisons have already taken the data variance into account as a kind of adjustment in the computation, by not purely considering the mean value comparison.

Often when the data distribution is not normal, it does not look like a bell shape in other words, the median and inter-quartile range (IQR) instead of the mean and standard deviation (SD) are used to describe the data. For non-normal distribution of data, the IQR in the tables and figures below should therefore be focused instead of the SD when looking at the data dispersion. Both are measures of data dispersion. SD measures the dispersion from distributions following normality assumption; while IQR measures the ‘non-normal’ one. The interpretation is similar. Larger value indicates larger dispersion. However, the values of estimates of the two statistics cannot be compared directly. In other words, SD = 1.9 cannot be compared with IQR = 1.84. The comparison will be meaningful when the statistics of different groups are
compared. For example, the data dispersion of group A, which has SD = 2.5 > group B, which has SD = 1.6.

To graphically represent the IQR data dispersion in the figures below, the top of the box plot represents the 75th percentile, the bottom of the box plot represents the 25th percentile, and the line in the middle represents the 50th percentile. The whiskers, the lines that extend from the top and bottom of the box plot, represent the highest and lowest values that are not outliers or extreme values. Percentile is a kind of order statistics. The 25th percentile means the cutoff value that 25% of data have value below the cutoff; this is also named as cutoff lower quartile or Q1. Precisely, the 25% lower quartile cutoff fixes that 25% of the total number of (ranked) subjects are below the cutoffs. Similarly, 75th percentile means the cutoff value that 75% of data have value below the cutoff; this is also named as cutoff upper quartile or Q3. The 50th percentile means the cutoff value that 50% of data have value below the cutoff; it is equivalent to the 'median'. The difference between Q1 and Q3 (Q3-Q1) is the inter-quartile range that can be treated as a measure of the data dispersion. Larger IQR represents large variance in data and vice versa. Graphically, if the lower part of the box plot is taller than the upper part, then the data are relatively more dispersed in the 'lower part' but less dispersed in the upper part. For example, the Maths scores of 10 students in a class have the following distribution -18, 22, 35, 40, 45, 48, 49, 50, 50, 52. The lower 25% data (i.e. score 18, 22, 35) have a higher dispersion than the upper 75% (i.e. 50, 50, 52). Outliers, values that are between 1.5 and 3 times the interquartile range, and extreme values, values that are more than 3 times the interquartile range, are represented by circles beyond the whiskers.
Table 37 – Results of One-way ANOVA and Post Hoc Multiple Comparisons among Three Feedback Treatments on Students’ Past Tense Used Ratio

<table>
<thead>
<tr>
<th>Treatment group</th>
<th>Descriptive statistics</th>
<th>One-way ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Control (CNT)</td>
<td>0.198</td>
<td>0.189</td>
</tr>
<tr>
<td>Normal Recast (NR)</td>
<td>0.328</td>
<td>0.174</td>
</tr>
<tr>
<td>Corrective Recast (CR)</td>
<td>0.340</td>
<td>0.140</td>
</tr>
</tbody>
</table>

Note: IQR represents the inter-quartile range; C denotes that the mean ratio is statistically significantly higher than the control group by post hoc comparison Scheffe test at 0.05 level of significance.

As the data did not follow the normality assumption, in other words not distributed normally or statistically significantly departing from the normal distribution, the mean and SD cannot describe the pattern of the distribution accurately and the t-test may produce invalid result. Therefore, the non-parametric test, Mann-Whitney U Test, was used instead. The variables measured across the two feedback sessions were more how-many-time variables than happen-or-not. How-many-time is of a continuous and scale nature; whereas happen-or-not is of a
dichotomous and nominal data nature. Same as the above case with the ANOVA test in table 37 and figure 11, interpretation should mainly be based on the statistical test results. The figures merely serve as assistance, for the purpose of description.

Table 38—Mann-Whitney U Test on Variables across the Two Feedback Groups

<table>
<thead>
<tr>
<th>Treatment group</th>
<th>1. Normal Recast feedback (NR) (n=29)</th>
<th>2. Corrective Recast feedback (CR) (n=29)</th>
<th>Z</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Errors</td>
<td>Mean 11.93</td>
<td>11.55</td>
<td>-0.101</td>
<td>0.919</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation 4.50</td>
<td>3.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median 12.00</td>
<td>12.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IQR 6.00</td>
<td>4.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RecastAttempts</td>
<td>Mean 11.79</td>
<td>11.10</td>
<td>-0.391</td>
<td>0.696</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation 4.40</td>
<td>3.89</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median 12.00</td>
<td>12.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IQR 6.00</td>
<td>5.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recasts</td>
<td>Mean 11.79</td>
<td>10.10</td>
<td>-1.396</td>
<td>0.163</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation 4.40</td>
<td>3.63</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median 12.00</td>
<td>10.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IQR 6.00</td>
<td>5.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SuccessUptakeReformulation</td>
<td>Mean 8.07</td>
<td>7.69</td>
<td>-0.359</td>
<td>0.72</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation 5.54</td>
<td>4.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median 9.00</td>
<td>8.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IQR 9.00</td>
<td>5.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UnrecastErrors</td>
<td>Mean 0.14</td>
<td>0.45</td>
<td>-1.721</td>
<td>0.085</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation 0.35</td>
<td>0.74</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median 0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IQR 0.00</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UptakeNoReformulation</td>
<td>Mean 0.17</td>
<td>0.41</td>
<td>-0.178</td>
<td>0.859</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation 0.38</td>
<td>1.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median 0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IQR 0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SuccessUptakeHalfway</td>
<td>Mean --</td>
<td>1.00</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation --</td>
<td>1.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median --</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IQR --</td>
<td>2.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NoUptake</td>
<td>Mean 2.90</td>
<td>1.76</td>
<td>-0.832</td>
<td>0.406</td>
</tr>
<tr>
<td></td>
<td>Standard Deviation 4.56</td>
<td>2.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Median 1.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>IQR 4.00</td>
<td>2.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IncorrectUptake</td>
<td>Mean</td>
<td>0.66</td>
<td>0.24</td>
<td>-2.447</td>
</tr>
<tr>
<td>----------------</td>
<td>---------</td>
<td>------</td>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.86</td>
<td>0.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>0.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IQR</td>
<td>1.00</td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: IQR represents the inter-quartile range. * < 0.05

Considering the feature of the data distribution, the non-parametric Mann-Whitney U Test was applied to test if there was any statistical difference between the normal recast and corrective recast groups among the 8 aspects in the feedback sessions. However, the Success Uptake Halfway variable was an exception, because only the corrective recast group had data on students' uptake of the use of past tense after the question prompt and before the subsequent corrective recast. The results showed that statistically significant difference between the normal and corrective recast groups was found only in the aspect of Incorrect Uptake. For other measures which showed slight visual differences in the figures below, they were actually not statistically significant according to the test.

Box plots were used below to demonstrate the distribution of data of the two groups on the 9 variables. Outliers and extreme values were also identified in the plots. First of all, in figure 12, the circle indicates the outlier which had values ranged from median +/- 1.5*IQR to median +/- 3*IQR. IQR represents the inter-quartile range. Case 6 in the corrective recast group had identity code CR3.21.SR in the dataset.
Figure 12—Distribution of ‘Errors’, ‘Recast Attempts’, ‘Recasts’, ‘Success Uptake Reformulation’ of NR and CR Groups

In figure 13 below, the measures of the variables did not carry any information as they barely had data dispersion. There were too many zero data. As shown in the medians, some lower quartiles were all zero. The figure however offers the information that a few individuals performed extraordinarily. The circles and crosses indicate the outlier and extreme values in the data distribution. Outliers had values ranged from median +/- 1.5*IQR to median +/- 3*IQR and extreme values had values higher than 3*IQR. IQR represents the inter-quartile range. Cases 31, 34, 36, 41 and 44 in the normal recast group had identity codes NR4.6, NR3.12, NR3.23.SR, NR3.41 and NR3.47.SR respectively in the dataset. Cases 4, 5, 14, 15, 17, 18 and 24 had identity codes CR2.18, CR3.20, CR1.55.SR, CR3.57.SR, CR3.59, CR4.62.SR and CR1.79 respectively in the dataset.
Concerning the following last single case for the corrective recast group, no comparison can be made because only this group had data available. The figure shows an equal distribution of the upper and lower quartile. Furthermore, after checking with the normality test 'Kolmogorov Smirnov test', the distribution had no statistically significant departure from the normal distribution. Therefore, the data of the variable presented in figure 14 can be interpreted as normally distributed.
Figure 14—Distribution of ‘Success Uptake Halfway’ of CR Group

In view of all the outlier and extreme cases shown in figures 11-14, it may be worth delving into what happened to students in the study in a qualitative way. The next chapter, chapter 6, reports the inner thoughts of a subset of students at the moment of receiving the different feedback types in the study through the use of stimulated recall interviews.

5.10 Chapter Summary / Conclusion

This chapter has analysed the experimental data via the different relevant statistical tests, to verify the different variables’ effects as predicted by the present study’s hypotheses. In sum, the two experimental feedback types (corrective recast and normal recast) facilitated learners’ use of past tense more significantly than the control feedback type (content-only) in their immediate posttests’ spoken narratives. This seems to confirm hypotheses 1 and 2. However, only the normal recast feedback was shown to have more significantly promoted learners’ use of past tense than the
control feedback type in their delayed posttests' spoken narratives. Hypothesis 3 was therefore confirmed but hypothesis 4 was rejected. Moreover, both hypotheses 5 and 6 were refuted because the corrective recast and normal recast feedback types did not appear to have been significantly different from each other in the immediate and delayed posttests.

For data beyond testing the research hypotheses, students from the different feedback groups have been shown using both accurate and developing form of past tense across time. Students’ lists of verbs presented also reveal the insight of their use of other forms both counted and not counted in the dataset. The ANOVA analyses of all these different forms demonstrate specifically how students’ spontaneous use of the different forms in narrative speech differed across time and feedback groups. Last but not least, only the variable of Incorrect Uptake among the 8 comparable variables all together was found significantly different between the normal and corrective recast groups. This suggests that the two feedback groups were not too different from each other in terms of the number of feedback turns given, students’ uptake of the corrections, or no uptake at all for example.
Chapter 6 Qualitative Data Analysis

6.1 Introduction

Contrary to the quantitative results of the present study, the more facilitative effect of enhanced recasts than unenhanced recasts has been documented by some past studies discussed in chapter 2. An in-depth qualitative investigation may be helpful in revealing what possibly caused learners’ not benefiting more significantly from corrective recasts than normal recasts. The qualitative investigation made use of stimulated recalls to motivate learners to recall their states of mind at the time of receiving the different feedbacks (normal recast, corrective recast, content-only) and using past tense in their spoken narratives. This chapter analyses learners’ different cognitive encounters.

For clarification, the qualitative investigation focuses on providing a perceptive angle of viewing the study’s results, instead of illustrating learners’ performance change. Overall speaking, the qualitative findings exemplified that learners’ awareness of the different feedbacks and use of past tense were basically affected by the task demand, the task modality, learners’ cognitive constraints, learners’ language development at the time of data collection, and the effects of recasts and learners’ uptakes.

6.2 The Qualitative Study Rationale

The qualitative analysis attempted to broaden the scope of the present study by looking in-depth into learners’ states of mind during their different feedback sessions. This was accomplished by continuously asking the central question of “What were you thinking at that time?” (Gass & Mackey, 2000), and letting learners freely verbalize their awareness of their mistakes, use of past tense, recasts and phenomena
occurred during the feedback sessions. This may serve to unveil the reasons behind some of learners' quantitative performance findings. For analyzing the stimulated recall data, the approach of thematic analysis was used to reveal issues that became evident from the dataset and that had not been initially considered in the study's hypotheses. Learners' states of mind data analysed in this chapter were selected and organized according to the recurring issues arising from learners' stimulated recall data. There were no clear differences between the normal and corrective recast groups' stimulated recall comments, as the recurring issues arisen were the same.

Excerpts of the stimulated recall data are demonstrated in appendix XIII. Last but not least, the chapter will end with a summary of the major qualitative findings.

During the stimulated recall session, both the researcher and students held the initiative to pause the feedback session video and responded to the selected excerpts. The researcher particularly paused the video when students exhibited: 1) pauses, 2) stammer, 3) repetition of the same words, 4) verbless sentences from L1 transfer, 5) incorrect uptake of normal or corrective recasts, 6) successful uptake of normal or corrective recasts, 7) initiation to comment, 8) continuous mistakes after attempts of normal or corrective recasts, 9) wrong, confusing or incomplete meaning or use of vocabulary, 10) inconsistent or consistent use of past tense, 11) laughs, or 12) unclear pronunciation. The identity code of each stimulated recall learner is constructed as shown in table 39. Other types of coding used in the stimulated recall excerpts are as shown in table 40. Keys for the feedback session transcription are listed in appendix VIII.

<table>
<thead>
<tr>
<th>Code</th>
<th>Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NR1.1.SR</td>
<td>The first learner in the entire dataset (1.SR) who received normal recasts (NR), followed the first (NR1) order distribution of cartoon-strips (see appendix VI), and participated in stimulated recall (SR)</td>
</tr>
<tr>
<td>CR2.2.SR</td>
<td>The second learner in the entire dataset (2.SR) who received</td>
</tr>
</tbody>
</table>
corrective recasts (CR), followed the second (CR2) order distribution of cartoon-strips (see appendix VI), and participated in stimulated recall (SR)

| CNT3.3.SR         | The third learner in the entire dataset (3.SR) who received content-only feedback (CNT) as the control group, followed the third (CNT3) order distribution of cartoon-strips (see appendix VI), and participated in stimulated recall (SR) |

<table>
<thead>
<tr>
<th>Code</th>
<th>Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>Student’s response in the stimulated recall</td>
</tr>
<tr>
<td>R</td>
<td>Researcher’s question/response in the stimulated recall interview</td>
</tr>
<tr>
<td>Number</td>
<td>Line on the video clip of the treatment session</td>
</tr>
</tbody>
</table>

Table 40—Transcription Code in Stimulated Recall Excerpts

6.3 Task Demand

6.3.1 Sequential or Continuous Picture Narrative

This section first of all illustrates how students from the different feedback groups (corrective recasts, normal recasts, and content-only) spoke about their problem dealing with the sequential or continuous nature of the cartoon-strip narrative task. The different excerpts showed that having to mind what happened next in the story caused obstruction to students’ attention to other aspects.

NR1.10.SR (table 41 in appendix XIII) received normal recast on his/her non-past use of “is die”. Although s/he had successful uptake of the past form “was die”, s/he paused and stammered before and while saying the next sentence. The student reported that s/he noticed his/her continuous mistake of past tense. However, his/her recall revealed that his/her priority at that time was to construct the next meaning, by wasting no more time or concentration on his/her previous mistake. The student also said s/he was thinking of both the meaning and tense use of the next sentence; s/he may have therefore used “had” instead of “have”. However, the ambiguous difference
between “have” and “had”, which share the same vowel in speech, still confounded the listener before the in-depth inquiry via stimulated recall.

Having to worry about the following part of the story also made CNT1.11.SR (table 42 in appendix XIII) unable to finish his/her first meaning in line 24. When the student spoke about the sun children, he/she interrupted his/her own flow by indicating a false start of the meaning. He/she then started a new narrative of the Goddess of the East. The meaning about the sun children had thereby been left out. The student recalled that s/he was preoccupied with the following meaning at that moment; and that may have shifted his/her attention away from finishing the original meaning s/he wanted to say about the sun children. The foresight of planning the next immediate meaning in the sequential narrative task may have overshadowed the student’s thinking or organization of the current meaning, as well as his/her use of past tense.

Even when the student managed to complete the current meaning, s/he may not have been able to spare equal amount or enough attention to proceed to the next meaning without delays. CNT1.13.SR (table 43 in appendix XIII) paused for a significant while after his/her narrative of the suffering people. S/he recollected that s/he was stuck, because s/he was still pondering on the use of words to verbalize the next meaning. The sequential nature of the story narrative task, where students’ attention was expected to be on more than one meaning at a time to keep the continuous flow, may have overloaded CNT1.11.SR and CNT1.13.SR’s attention and allowed them to only convey meaning of either the current or the next story part or picture at that time. Because of the overloaded attention to constructing one meaning at a time, CNT1.11.SR and CNT1.13.SR may have become unaware of the use of past tense in “tell” and “do not have” respectively.
Being distracted by the following meaning was also shown in CNT4.17.SR’s (tables 44-46 in appendix XIII) three stimulated recalls. In lines 12, 24-26 and 39-40, the student kept repeating the same words. S/he informed that s/he was thinking about the next meaning while narrating the current meaning. Repetition of words may have been the student’s way of buying time for him/herself to shape the following meaning and avoid a broken flow. Narrating the current meaning and simultaneously maintaining the consecutive thought for the ongoing story may have retarded the student’s immediate progress to narrating the next meaning. Awareness of the use of past tense may have therefore been too much a multi-tasking burden for the student.

Connecting two consecutive pictures was also shown to have been the student’s primary concern. CR3.21.SR (table 47 in appendix XIII) did not notice his/her mistake with not using past tense for “eat” from the hint given in the question prompt. S/he tried to continue with his/her unfinished meaning instead. S/he was then able to take up the mistake after the subsequent corrective recast. However, his/her uptake could have been purely a repetition or not, because the recall unveiled that his/her focus at that time was on associating the two consecutive pictures. This focus on the following meaning had been prolonged to his/her next narrative (line 9) of his/her originally incomplete meaning.

Staying engrossed in the following meaning had also affected the student’s proper use of language and thereby the expressivity of his/her original meaning. CNT1.22.SR’s (table 48 in appendix XIII) intention to communicate the meaning accurately was reported to have been impeded by his/her preparing against breaking up the narrative flow and story development. The meaning of “in the sky” had not been conveyed clearly from the student’s mind which was swamped with the next
meaning gap needed to be filled at that time. His/her misuse of tense may have also been caused by the apprehension over expressing the upcoming meaning.

The non-fragmented nature of the cartoon-strip narrative task prompted the student to assiduously move forward the storyline to speak about the next plot. As recalled by NR3.23.SR (table 49 in appendix XIII), thinking of the character’s next action or response to the sea of fire was his/her centre of attention at that time. His/her successful uptake of the normal recast on past tense could have been an attentive uptake or purely a repetition; because of the student’s recollection of minding the next immediate story plot and nothing about his/her mistake.

The urge to move forward the storyline sequentially had also cast difficulties on the student from the very beginning. Before delivering any story plot, CNT2.39.SR (table 50 in appendix XIII) already found it difficult to proceed with the narrative. S/he remembered that she was stranded by the prospect of linking the 12 sequential pictures to form a story narrative from start to finish. The constant transition from one story plot presented on each picture to another along the timeline had been the student’s concern of difficulty at that time. The pre-task planning did not seem to have been helpful to the student, in preparing him/her for the forthcoming task. The on-the-spot tension of narrating the present picture and connecting all the following pictures may have overwhelmed the student from the very start.

Interpreting and communicating the wrong or confusing meaning of a subsequent picture was shown to have been CNT2.40.SR’s (table 51 in appendix XIII) recourse to remedy his/her inability to channel the meaning of the next picture. S/he was uncertain about the following meaning to complete his/her original sentence. CNT2.40.SR admitted in his/her stimulated recall that s/he was actually unready to deliver the meaning. His/her narrative in line 15 may have therefore been a reckless
response to the push to continue with the sequential flow. His/her narrative in line 15 also seemed repetitive of that in line 13. The repetition was quite convincingly a heedless resort, just to manage the continuous narrative. The student's unawareness of using past tense in "tell" could have been caused by his/her quandary about the meaning coming up.

To avoid losing track of the meaning of the subsequent pictures, NR3.43.SR (tables 52-53 in appendix XIII) had been spending considerable effort on the next story plot trailed behind. S/he explicitly revealed that s/he had to keep an eye on what needed to be said next at the expense of the use of tense. The student was aware of his/her mistake as pointed by the normal recast, or even the use of past tense in the narrative. However, s/he could not retain the consistent use of past tense (line 39), possibly due to his/her priority over connecting to the following meaning.

Upholding an unbroken transition from the previous meaning to the next may have overridden the student's attention to his/her own mistake and use of past tense in the narrative. NR3.47.SR (table 54 in appendix XIII) did not show any verbal uptake of "gave" and continued with the subsequent meaning instead. There was no overt clue whether s/he noticed the normal recast correction, as s/he incessantly proceeded with the development of the story. The stimulated recall suggested that NR3.47.SR may have interpreted the normal recast as the researcher's mere expression of meaning, and therefore mainly engaged in bridging the researcher's meaning with his/her narrative of the next story plot.

Focusing on driving the storyline forward may not deprive the student of his/her awareness of the use of past tense. CR1.55.SR (table 55 in appendix XIII) recalled that s/he was aware of the hint given by the corrective recast, as well as bearing the motivation or readiness to undertake the subsequent narrative. However,
the student showed an inaccurate uptake of the form targeted. This indicated that surface uptake of the form may not have reflected what the student noticed at that time. The wrong uptake of “told” may have only been a slip of the tongue or a distraction from having to continue shortly.

Even when the student showed an accurate uptake of the form targeted, there was no guarantee that s/he was aware of his/her mistake. CR4.58.SR (table 56 in appendix XIII) had been hoping to compensate what s/he had not yet finished and express what s/he previously had in mind, but no clear sign of his/her awareness of using past tense for “take”. This was supported by the student’s completion of his/her originally unfinished pronunciation of “Children” or correction of his/her initially wrong pronunciation, as a response to the pre-corrective recast question prompt; his/her expansion of meaning from taking care of the children to taking care of the children and the Emperor; and most convincingly his/her recall that stretching the storyline was what s/he inclined towards at that time.

The pressing need to mind the following story plot could have still been haunting the student and affecting his/her use of past tense, regardless of his/her awareness of the recurring mistake and the consistent aid from feedbacks. NR2.60.SR (table 57 in appendix XIII) had clear uptake of the normal recast. S/he however confided that his/her inability to use grammar cautiously beforehand and avoid misusing past tense repeatedly was due to the concern of progressing the narrative.

An already-planned narrative of the forthcoming picture could have been lingering in the student’s head while s/he was narrating the current picture. Alternatively, the student could have been preoccupied with the next meaning before speaking about its preceding picture. CR4.62.SR (table 58 in appendix XIII) gave no sign of his/her awareness of the corrective recast on the target form. S/he was
nevertheless aware of the use of past tense, as shown in line 53. His/her use of “did not love” further suggested that s/he should have had no problem using the past auxiliary form “did”, which was being targeted by the corrective recast. The unclear uptake suggesting the student’s unawareness of the mistake may be affirmed by his/her recall that s/he was consumed with the meaning of the next picture at that moment. Regarding the unclear uptake, there was also a possibility that CR4.62.SR felt it was redundant to repeat the already-known form “did”. The drift of the next meaning was anyway the focus of the student.

Not only did the meaning of the subsequent picture bother the student, but also the means to translate the meaning of that picture. CR4.65.SR (table 59 in appendix XIII) showed obvious uptake of the corrective recast; and s/he was alert to the accurate use of vocabulary, as in line 4 where the student elaborated on the description of “arrows”. CR4.65.SR’s stimulated recall additionally informed that s/he was mindful of the use of vocabulary to channel the subsequent meaning to the listener. The student may have been aware of both his/her past tense mistake and the use of vocabulary in the next meaning, or mainly to the next batch of words to use.

The consistent and corrective effects of the feedbacks may have been attenuated by the challenge of sustaining the relentless course of the cartoon-strip narrative. NR1.69.SR (table 60 in appendix XIII) had always been aware of his/her recurrent mistake and the need to use past tense in the narrative; but s/he confessed that the momentum of using past tense throughout could not be maintained. The momentum was reported to have been interrupted by upholding another momentum of not breaking up the sequential storyline.

Being aware of the mistakes with past tense and subsequently extending the use of past tense to another verb may have been less a problem than pondering the
unready construction of the next meaning. NR3.74.SR (tables 61-62 in appendix XIII) was able to take up his/her mistake, and even generalize the use of past tense to a following verb in line 34. However, the student recollected that thinking what words to use to fill the next meaning gap required quite some endeavour and time at that moment. Therefore, it was doubtful if NR3.74.SR had spare attention enough for his/her thoughtful uptake of the tense mistakes, instead of just empty mimicking.

Contemplating a new sentence to convey the forthcoming meaning may have disrupted the student’s fulfilment of a completed uptake. CR2.80.SR (table 63 in appendix XIII) showed quick and successful uptake of the auxiliary form “did not”, but no uptake of the main verb “hurt” attached to it. This technically successful uptake but perhaps grammatically unsuccessful uptake may have been caused by the divided attention to depicting the picture trailed behind, as recalled by CR2.80.SR.

The simultaneous cognitive involvement in preparing for the current picture as well as the next picture had been explicitly stated by CR4.85.SR (table 64 in appendix XIII) as demanding. S/he clarified that the current picture itself was not tough to deal with. It was rather the dual task of linking and starting another new picture shortly after the current one which was challenging. The student’s surface uptake of “became” may have been pure mimicking, because s/he recalled being engaged in finding the appropriate means to clearly express the perhaps already-understood meaning of the current picture.

6.3.2 Task Familiarity

This section demonstrates another task challenge. Some students were aware of the feedbacks given to their tense problems in the narrative. However, they recalled
that the obstructive force against their consistent use of past tense was their unfamiliarity with the story narrative task.

6.3.2.1 Not Familiar with Story Narration

First of all, CR3.21.SR (table 65 in appendix XIII) had been using past tense quite consistently until s/he committed the mistake of using “is”. Although there was no evidence whether the student took up the corrective recast on “was”, his/her stimulated recall unveiled the hidden awareness of the mistake. CR3.21.SR seemed to attribute his/her inability to keep the consistent use of past tense for all the verbs used at that time to the novelty of the narrative task mode. Therefore, the student may mainly have had problem integrating the use of past tense with the narrative task mode, rather than problem using past tense in general.

NR1.69.SR (table 66 in appendix XIII) even spelled out in his/her stimulated recall that the narrative task mode had exacerbated his/her already immature use of tense in L2. S/he specified that his/her inexperience with story narratives included not being skillful enough to fit words into the construction of meaning promptly and spontaneously. This may have imposed extra barrier to NR1.69.SR’s confident use of tense in L2 at that time. The broken narrative in Line 10 may have been caused by the student’s inept way of handling story narrative. The subsequent mistake of not using past tense in “get” may have been the corollary.

The mismatch between the current experience and the student’s past experience with story narrative may have also led to his/her misuse of tense. NR3.74.SR (table 67 in appendix XIII) was aware of the use of past tense and used it correctly at the beginning of line 62. However, s/he did not use past tense for the two verbs in the last narrative of the entire story. The student defended in his/her recall
that using non-past tense for the narrative ending was legitimate in his/her past experience. This may have been why s/he naturally turned back to not using past tense for “leave”, regardless of the previous normal recast on “was” and his/her awareness of it. NR.3.74.SR’s ingrained experience in using non-past tense for the story ending may have also convinced him/her to deliberately follow what s/he knew before.

6.3.2.2 Familiar with Other Tasks

Instead of being familiar with story narrative, some students revealed in their stimulated recalls that they were adept at other types of tasks which had been given by teachers. The novelty of the sequential narrative task mode may have made CNT2.39.SR (table 68 in appendix XIII) want to withdraw from doing it initially. S/he clearly said that describing a single picture at one time was actually what s/he was comfortable with because of the regular practice at school; whereas describing a series of pictures all at one time would disconcert him/her and make him/her almost speechless at that time.

Unfamiliarity with the story narrative task mode may have caused uncertainty and therefore anxiety to the student about what s/he had just said. CNT2.46.SR (table 69 in appendix XIII) exhibited communication breakdown and was unable to deliver the meaning in the middle of his/her own narrative. The meaning disruption was recalled to be due to the student’s realization of his/her mistake with not using past tense and puzzle over the meaningfulness of his/her previous narrative. CNT2.46.SR attributed his/her agitation to the pressure of adapting to an unfamiliar task design with which s/he seldom encountered in daily schooling. As s/he indicated, doing individual presentations was their school routine on the other hand; and it would have helped him/her avoid the problem. 
Unfamiliarity with the tense use in historical story narrative and the less controlled nature of the story narrative task in obligating the student’s use of past tense may have loosened his/her diligence to use past tense throughout. CR4.62.SR (table 70 in appendix XIII) recalled that s/he was aware of the mistake; and attention to constructing the story meaning surmounted that to using past tense for conveying the meaning. The student pointedly mentioned the ease that s/he would have had with the fill-in-the-blank task, where s/he would have been restricted to use past tense effortlessly. This was because of its predetermined meaning not requiring his/her diverted attention to construct meaning from pictures, and its format with gap-filling blanks consistently prompting his/her use of past tense. CR4.62.SR’s unskillful handling of both the meaning construction and spontaneous tense use in the story narrative task may have disordered his/her persistent use of past tense.

6.3.3 Task Completion

6.3.3.1 Detail Coverage

This section reports cases where the task demand of pictorial narrative was suggested to have imposed a sense of mission to uphold task completeness on students. For instance, some students put their concentration on examining whether they have covered all the details of the current picture. CNT4.17.SR (table 71 in appendix XIII) used non-past tense for the irregular verbs “are”, “become” and “is”, and paused time after time. S/he reminisced that cross-checking what description had been left behind his/her previous narrative of that picture engaged him/her at that time. This may have been why the student paused and perhaps thought for a bit before embarking on narrating a new detail. Being cautious of not missing any points on a
single picture and thereby fulfilling task integrity may have also diffused CNT4.17.SR’s alert to the use of past tense in the narrative.

Connecting every detail to shape the picture and looking in retrospect if the details were deployed fully to convey the entire meaning of the picture were recalled as CNT2.44.SR’s (table 72 in appendix XIII) vexation while narrating the picture at that time. His/her use of linking words such as “and” and “so” in the narrative and pauses before as well as after the use of them suggested that CNT2.44.SR may have been carefully trying to incorporate every single detail to express the picture meaning coherently. His/her additional monitoring of the incorporation of details may have all together depleted his/her attention to use past tense in the narrative.

Simply conveying the main meaning of the picture may not have been what CNT2.46.SR (table 73 in appendix XIII) aspired to do at that time. The picture narrative in line 8 was cued by the pre-task Chinese summary as only describing the situation that Chang Or and Archer God went to the earth and felt the heat in person. The student however elaborated the narrative more by also considering the background details about the surrounding people and animals in the picture. Such ambition and carefulness of ensuring sufficient detail coverage of the narrative may have already occupied CNT2.46.SR’s cognitive load, leaving the use of past tense aside at that moment.

Adding background information to the picture may have on the other hand been CR4.58.SR’s (table 74 in appendix XIII) way to avoid being imprudent. The picture at that time only showed the improved conditions of the plants and animals; but the student added a transition sentence from the previous incident of Archer God’s shooting the nine suns to the present picture. The inclusion of the transition sentence may have been motivated by CR4.58.SR’s meticulous mentality of safeguarding the
appropriate amount of details in the narrative. The student concentrated on keeping track of the description coverage of that picture. Regardless of such an endeavour, CR4.58.SR was able to retain the use of past tense. This may have been the effect of the persistent corrective recast supplies beforehand.

Stressing on the perfection of delivering the meaning of the present picture in entirety was also NR1.69.SR’s (table 75 in appendix XIII) focus at the time of narration. The difference with NR1.69.SR was that s/he employed not breaking up the description as the means to achieve the perfection of a complete description of the picture. The narrative in line 62 was rather lengthy and involved multiple actions. Striving to maintain a smooth-spoken narrative of these actions was reported to have helped him/her realize a complete narrative of that picture. His/her perception of the need to well-speak the complete meaning may or may not have overshadowed his/her awareness of his/her own uptake of the tense mistake.

Apprehension over insufficient detail coverage even extended to the student’s narrative of the very last picture. NR3.74.SR (table 76 in appendix XIII) had actually covered all the necessary details of that last picture to convey the meaning and call the story narrative to an end. S/he however was still unsure if there were some more details left to be said, regardless of the fact that the story had no room for extension in the very end. Such paranoid thinking may or may not have shrouded his/her attentive uptake of the tense mistake from the normal recast.

Being meticulous about the precise compositions of the picture was recollected by NR4.87.SR (table 77 in appendix XIII) as his/her simultaneous thinking while s/he was narrating the picture at that time. The accurate narration of the characters, the number of the sun children, and the order of happenings about these characters on that picture suggested the students’ exhaustive view of the picture.
His/her detail-oriented mindset at that time did not seem to have overtaken his/her discipline of using past tense for that picture's narrative. This may have been the effect of the consistent normal recasts throughout.

6.3.3.2 Ending Story

Some other students prioritized closing the entire story and presenting it intact, to fulfil task completeness. This section analyses students’ verbalization of their concern over finishing the story narrative task.

First of all, NR1.10.SR (table 78 in appendix XIII) did not recall being aware of his/her uptake of past tense, but explicitly said s/he was cogitating an approach to bring the narrative to an end. S/he paused in line 50 and may have been thinking through the ending of the story until the researcher’s normal recast intervention. After the recast, the student showed clear verbal uptake of the mistake as well as clear narrative of the storyline ending. His/her focus on completing the whole story may have contributed to the articulate ending and thereby a possible absent uptake of “told”.

The advanced thinking of the next and final sentence to conclude the story narrative seemed to have disrupted the flow of CNT4.17.SR’s (table 79 in appendix XIII) current narrative and expression of meaning. His/her broken, hesitant and inaudible narrative in line 39 may have been the result of deploying most of his/her attention to think ahead how to communicate the story ending and signal the end of the task to the listener. The achievement of task completion may have overcome CNT4.17.SR’s understanding that the essence of communication is often based on the process of meaning elucidation.
The rush to finish the story may have even led to the student’s resort to a wrong narrative of meaning. CNT2.40.SR (table 80 in appendix XIII) recalled that s/he was eager to round off the narrative task, and motivation to achieve such goal may have been stronger than that to retain the accurate meaning in the utmost priority. The wrong meaning that all of the sun children were killed may have therefore been presented. Pursuing the sense of mere task attainment seemed to have outweighed CNT2.40.SR’s incentive to pursue effective message transmission and correct time reference.

Polishing the story ending was reported to be NR3.43.SR’s (table 81 in appendix XIII) primary concern after finishing describing the key meaning of the last picture. S/he had already delivered the significant action that Archer God’s wife flew to the moon after taking all the medicine of long life. However, in order to close the story beautifully, s/he may have wanted to refine the language of this key meaning or include the narrative of Archer God’s emotion about his wife’s action as shown on the picture. The kind of perfection or task completeness that NR3.43.SR sought at that time was said to be pertaining to bettering the quality of the ending, either in terms of language delivery or content enrichment. This higher aim may have obscured the corrective effect of the previous normal recast and his/her subsequent uptake of it.

CNT2.46.SR (table 82 in appendix XIII) on the other hand clearly showed his/her problem delivering language but seemed to have approached the content properly, when dealing with the ending of the story narrative. The student aimed at calling his/her narrative to an end initially, but s/he was not yet satisfied with the completeness of it and was unable to refine it on the spot. His/her search for completeness may have been pertained to his/her language delivery imperfection, "just one..."; while the content words "just one" remained to provide hint on the key
meaning to the listener. His/her later confirmation of the researcher’s language plus meaning intervention suggested an echo to the students’ thinking at that time.

Terminating the story narrative without delay was reminisced as what CR4.62.SR (table 83 in appendix XIII) determined to achieve at that time. His/her sole disposition to resume his/her original narrative after the corrective recast intervention so as to complete the task may have driven his/her concentration away from taking up and being aware of the feedback and mistake. Alternatively, CR4.62.SR may have noticed the feedback but was so eager to complete the story that s/he may have felt a verbal uptake was redundant. S/he may have felt it was more direct and important to immediately bridge the incomplete meaning of the researcher with the remaining meaning, to wrap up the entire narrative.

Not taking up the mistake and mistaking the feedback as targeting something other than past tense may have been triggered by CR4.65.SR’s (table 84 in appendix XIII) emphasis on meaning communication. S/he particularly verbalized that the purpose of stressing on meaning communication at that time was to bring the story to an end. The value of content appeared to have been greater than that of language and acoustic delivery in favouring task accomplishment from CR4.65.SR’s point of view. His/her standpoint of using content to fulfil task completeness at that moment was most apparent when s/he successively stated the object matter (line 69) of his/her original narrative (line 67) after the question prompt interruption.

The final meaning of not coming back to the Heaven had been repeated by CR4.85.SR (table 85 in appendix XIII) throughout the feedback episode. The insistence upon bringing out this final meaning may have been engendered by the student’s obsession with settling the story narrative task at that time. Accordingly, his/her surface uptake of “did” may have been a mechanical mimicking. The narrative
of this final meaning as a means leading to task completeness may have been lingering in CR4.8S.SR’s mind throughout. This could have been why the final meaning kept reappearing immediately after each of the researcher’s responses.

6.3.4 Task Organization

6.3.4.1 Link up Ideas

Apart from the task demands of maintaining the sequential flow of narrative, task unfamiliarity, and upholding sense of task completeness, the task demand of narrative organization had also been recalled by some students as engaging them. For instance, CNT1.11.SR (table 86 in appendix XIII) pointed out that his/her narrative lacked an organized framework where an introduction was expected to provide background information. The student may have been considering an audience-oriented way to present the story by orienting the listener to the story setting before commencing the story content. CNT1.11.SR showed hesitation and may or may not have used past tense for “have”, with his/her mind engrossed in structuring the story in a more well-planned manner. CNT2.46.SR (table 87 in appendix XIII) even carried the thought of a well-planned narrative by means of an introduction to the middle of the narrative. S/he wanted to build an introduction specially for the individual picture at that time, perhaps to orient the listener better to the picture meaning. The student may have therefore exhibited broken narrative; and s/he may or may not have used past tense for “prepare”.

Other students showed mindfulness of the task’s organization by adopting a top-down way of thinking. In other words, they approached the narrative from a general scope down to a more detailed one. Both CNT4.17.SR (table 88 in appendix XIII) and CNT2.40.SR (table 89 in appendix XIII) may have traced back to the pre-
task Chinese summaries to consult or recapture the overall meaning or structure of the story, in order to better advise his/her organized narrative of the present picture. This may have led to their broken narrative, uncertainty about the meaning and vocabulary, and misuse of tense.

Assembling different components to form meaningful messages was also stated by some students as their inner thoughts of task organization while narrating. There were different elements on a single picture which worked together to provide clue to the meaning. As in the cases of CNT4.17.SR (table 88), CR2.37.SR (table 90), CR4.65.SR (table 91), and CNT4.77.SR (table 92) in appendix XIII, efforts which contributed to students’ ability to relate the elements and ultimately form meaning included organizing the order of happenings of the different elements and the relationship between the elements. These efforts may have consumed the students and left them nearly no room to note the use of past tense (CNT4.17.SR), the coherent narrative of meaning (CNT4.17.SR, CR4.65.SR), and their L1 transfer of legitimate verbless grammar in Cantonese (CNT4.77.SR). Nevertheless, CR2.37.SR managed to use past tense consistently and narrate meaning coherently, regardless of his/her preoccupation with associating the different elements in the picture.

Use of linking words may have been another way of fulfilling task organization in some students’ minds. CNT2.39.SR (table 93) and CR3.57.SR (table 94) in appendix XIII recalled having directed their attention to making their narrative sentences transit well from one to another and driving the development of story plots to proceed logically. CNT2.39.SR appeared to have been debating whether to use the cohesive device “so that” or “at last” to bring out the consequences of Archer God’s killing the sun children and leaving one sun, from the ineffectiveness of the arrows and the attitude of the sun children. His/her recollection of that narrative moment
suggested that connecting sentences, to shift from the causes to the outcomes to promote meaningfulness, was his/her prime focus. The clear sign of not using past tense in “leave” may have been resulted. CR3.57.SR also seemed to have been indecisive about the use of transitional time clause “when...” to project the link between the previous action of seeing the sea of fire and the next immediate action. This could be affirmed by the student’s stammer of the when-clause and his/her recall of relating the current narrative to the preceding one. His/her use of past tense had not been affected by channeling attention to the transitional means though.

The following four students shared the common experience of repeating the same or similar words or phrases in their narrative and showing concern of coordinating their thoughts while narrating. First of all, NR3.23.SR (table 95 in appendix XIII) was switching the subject and object nouns around the same verb “told”. S/he confided that s/he lost sanity of thinking and felt his/her logical narrative may have therefore been interfered. Although the use of past tense had been maintained, the meaningfulness was impaired as admitted by the student. His/her organization of comprehensible meaning was what NR3.23.SR disappointed with at that time. CR2.38.SR (table 96 in appendix XIII) was also upset about his/her arrangement of meaning during the narrative. The repetitiveness of some words and inconsistent use of past tense may have been caused by the realization of his/her being disoriented at that time. The verbal uptake of “became” may or may not have been heeded. The repetitiveness of some words may have also been the outcome of CNT2.44.SR (table 97 in appendix XIII) and CR2.80.SR’s (table 98 in appendix XIII) being immersed in formulating their thoughts. The difference was that the latter did not seem to have problem using past tense as a result.
6.3.4.2 Complete Sentence or Meaning

Persistence on completing their original narrative sentences had been given precedence by some students over other considerations in their minds. Not allowing feedback turns to break up their original meanings suggested that they laid much emphasis on maintaining the well organization of their narratives. The following students, NR1.10.SR, NR2.34.SR, NR3.47.SR, CR1.55.SR, NR2.61.SR, and NR3.74.SR, commonly had their attention centred on prolonging their unfinished initial meanings.

NR1.10.SR (tables 99-100 in appendix XIII) showed clear successful uptake of “did not work”, but s/he only recalled it as a fast repetition. This made it doubtful if his/her uptake reflected the student’s noticing the mistake. The further reason that NR1.10.SR spelled out why s/he had to quickly repeat the target form of the normal recast suggested his/her uptake may have been a futile one. S/he mainly wanted to be done with the repetition to save time to preserve and finish his/her previous meaning before it was forgotten. NR1.10.SR kept being preoccupied with the narrative s/he had not settled before the normal recasts in later parts of his/her feedback session. Line 39 and its corresponding stimulated recall indicated the student was unaware of the corrected form because of his/her prudence to keeping both his/her and the researcher’s narrative organized with full meaning. Line 42 and its corresponding recall did not evidently show the student’s awareness of the mistake, but his/her focus at that time remained as well organizing his/her initially broken meaning.

Similar to NR1.10.SR, NR2.34.SR (table 101 in appendix XIII) prioritized making up for his/her unfinished narrative to formulate meaningfulness, and may have therefore missed taking up the normal recast accurately. NR3.47.SR (table 102 in appendix XIII) and NR2.61.SR (table 103 in appendix XIII) exhibited clear verbal
uptake of the normal recast; but their immediate meaningful extension of narratives and recalls of primarily heeding the continuation of their broken meaning suggested uncertainty in whether they had taken up the tense mistakes. The case of CR1.55.SR (table 104 in appendix XIII) was different in the sense that s/he did not seem to have dwelled on continuing his/her previous unfinished narrative, but further elaborated on his/her previous finished narrative. The addition of concrete examples illustrating his/her previous finished narrative may have been CR1.55.SR’s way to well coordinate his/her thoughts.

NR2.61.SR (table 105) and NR3.74.SR’s (table 106) excerpts in appendix XIII also demonstrated students’ principal interest in formulating unbroken and meaningful narratives. The distinguishing cases that needed to be pointed out were NR2.61.SR’s unobstructed awareness of the tense mistake and NR3.74.SR’s disrupted tense consistency. From line 43 to 45, NR2.61.SR articulated that s/he noted his/her tense mistake, but the zero verbal uptake resulted was due to his/her attention to mending his/her initially broken meaning. S/he also revealed that s/he took priority of meaning completion so as to facilitate effective communication. For NR3.74.SR, s/he had been using past tense until one incidence in the excerpt. S/he did not verbalize his/her attentive uptake or not at that time in his/her recollection, though clear verbal uptake was shown during the feedback session. The same focus of organizing meaningful communication may have disrupted his/her persistent use of past tense and led to his/her blind verbal uptake.

Apart from continuing previously unfinished meaning to attain better organization of ideas, some students regarded repeating what s/he had once said but missed saying it later as fulfilling meaningful organization of ideas. For example, NR1.10.SR (table 107 in appendix XIII) recalled his/her suspicion over failing to
present a complete meaning. S/he may have meant the lack of the subject “God of the East” in his/her uptake, which had been spoken by himself/herself before as well as the researcher in the normal recast. This fear of conveying disorganized meaning may or may not have overtaken the student’s awareness of the tense correction.

CNT2.46.SR (table 108 in appendix XIII) even explicitly reported his/her wonder over repeating the part that s/he had narrated before the researcher’s intervention. His/her consideration of integrating an already-said part with his/her uptake of an incomplete meaning from the researcher suggested his/her concern for meaningful organization. This may have overshadowed his/her use of past tense.

Last but not least, CNT2.46.SR (table 109 in appendix XIII) overwhelmingly prioritized meaningful organization over the use of an accurate word. S/he admitted that s/he had doubt on the appropriateness of using “cheat” to convey the meaning that the sun children were laughing at Archer God. However, s/he forsook changing the word or finding another more suitable word at that time, just to avoid fragmenting the organization of meaning.

6.4 Speaking Modality

The speaking modality is different from the writing modality in terms of the limited time allowance for conceptual construction, unbroken flow requirement for effective communication, simultaneous planning at the moment of speaking, reliance on pronunciation as the medium for meaning transmission, and the fleeting existence of the spoken message (Cameron, 2001). Some students in the stimulated recall subset pointed out their encounters with these differences.

Both NR2.34.SR (table 110) and CR4.85.SR (table 111) in appendix XIII recalled their awareness of the use of past tense. However, they clarified in retrospect
that they did not really have problem using tense at that time, regardless of their overt mistake with past tense. NR2.34.SR revealed that s/he was unaware of his/her wrong actual use of tense at the moment of speaking. This may have been why s/he used “lead” both before and after the normal recast. CR4.85.SR explicitly raised the fact that writing and speaking are different in terms of speakers’ unawareness of their messages after the moment of speaking has passed. Mispronunciation of the tense form was recalled as the underlying mistake instead, which CR4.85.SR claimed not noticing it at the moment of speaking. Some students had also spoken about the issue of pronunciation, which will be analysed in a subsection here.

6.4.1 Insufficient Time

The live and on-the-spot performance nature of speaking did not appear to have allowed time for students to think more deeply before conducting the narrative of each part or picture. CR3.21.SR (table 112), CNT2.40.SR (table 113), CR4.58.SR (table 114), and NR1.69.SR (table 115) in appendix XIII shared the same recall of not having enough time for furthering their cognitive endeavour to facilitate their current narratives. First of all, CR3.21.SR left his/her narrative at that time rather brief, “…have clear a thing, one thing”, when in fact the picture illustrated what exactly the Goddess of the West wanted to say about taking the two medicine of long life. S/he recollected having struggled to elaborate his/her previous narrative. The obstructive force had been insufficient time for thinking how to elaborate. CNT2.40.SR was even unable to construct a complete meaning for his/her narrative, because of time deficiency during the speaking performance. Time deficiency may have also affected his/her use of past tense. CR4.58.SR attributed his/her tense mistake to his/her recurrent mispronunciation, and also lack of time for him/her to carefully mind his/her
actual pronunciation before speaking again in a split second. Last but not least, NR1.69.SR resorted to stammers and repetition of words already used, because of no rehearsal time to search for new and suitable words on the spot.

Some other students rather followed what they had already known to be wrong than amending it due to the lack of time to do so. CNT1.13.SR (table 116 in appendix XIII) recalled knowing how to describe the people on the picture, but revealed that s/he deliberately did not include it originally because of limited time allowance. His/her use of past tense may have thereby been compromised. Apart from leaving content inadequately described, surrendering grammar accuracy was also shown to be students’ recourse to sustain their narratives within scarce time. CNT2.39.SR (table 117 in appendix XIII) noticed his/her problem using grammar and thereby the confusing meaning, which made him/her anxious. S/he however did not attempt to make changes because of no time to ponder on the passing problems. Moreover, his/her nervousness continued to affect CNR2.39.SR’s clear mind for his/her subsequent narrative. CNR2.40.SR (table 118 in appendix XIII) showed consistent use of past tense in the excerpt. Nevertheless, s/he confided that the consistency did not last in other occasions, which s/he was aware of, because s/he was short of time to pay attention to keeping that. CR1.55.SR (table 119 in appendix XIII) also encountered problem maintaining tense consistency and was aware of that at that time; but s/he gave up going back to monitor and change the mistake because of little time. This was a similar problem for CR4.85.SR (table 120 in appendix XIII), who was reminded of the mistake and quickly corrected it halfway through the corrective recast attempted. S/he forsook correcting it initially because of the short window of time before communicating the subsequent meaning.
Besides the speaking modality itself imposing time limitation on speakers, the immediate audience of the speaking modality also projected an impression of pressing time to speakers. CNT1.33.SR (tables 121-122 in appendix XIII) had been using past tense accurately. However, s/he faced problem expressing meaning of the two pictures. In the beginning, CNT1.33.SR was striving to conceptualize the picture and then channel it into speech. However, s/he failed to do so because of inadequate time for delving into the use of language beforehand. S/he further unveiled the reason causing time pressure, which was fearing the listener would have his/her interest fallen off if time were taken longer than expected. Later on, CNT1.33.SR experienced meaning misinterpretation and s/he was aware of that. However, s/he did not attempt to make any changes at that time, because of the pressure of keeping the listener on the same narrative meaning for long.

6.4.2 Ongoing Flow

Time pressure in speaking was found induced by the commitment to adhere to the continuous demand of the cartoon-strip narrative in the study. Students like CNT1.11.SR (table 123) and CR4.58.SR (table 124) in appendix XIII expressed their obligation to keep up with the ongoing flow of the narrative. With the absorption in gluing the different story plots on individual pictures, CNT1.11.SR still noticed his/her missing an auxiliary verb before the main verb “listen to” in the narrative. However, s/he purposely did not make the amendment, to avoid interrupting the succession. CR4.58.SR also recalled that the task pushed him/her to go forward with the narrative. S/he may or may not have therefore attended to the corrective recast, though his/her successful uptake was shown.
Same as CNT1.11.SR’s case but more recurrent, NR3.74.SR (tables 125-126 in appendix XIII) had been noticing his/her problem interpreting and expressing meaning in his/her narrative. However, s/he chose not to put forward any changes at that time because of the pressure of continuing the communication flow. If not, his/her speech may have been hampered by silence and broken messages. Although NR3.74.SR cared about securing the ongoing flow, s/he did not seem to have forgotten the use of past tense. This can be shown from his/her successful uptake and subsequent use of past tense in other verbs.

Some students were on the other hand distressed by finding a way to support the continuous course of the speaking task. CNT1.13.SR (table 127) and CNT4.17.SR (table 128) in appendix XIII had been stammering through their narratives and repeating words to fill in the gap of silence while speaking. They recalled figuring out a means to fulfil the continuity of their narratives at that time. This may have been why CNT1.13.SR exhibited his/her L1 transfer of missing a verb, stammers, repetition of the same words, and no use of past tense. For CNT4.17.SR, s/he may have as a result been disorganized in arranging the order of happenings with the different characters and inconsistent in using past tense.

Instead of pondering a method to continue the flow, CNT4.17.SR (table 129) and NR3.74.SR (table 130) in appendix XIII concentrated on using their own ways to keep the communication flow. CNT4.17.SR made use of appropriate words to drive the communication flow going; and NR3.74.SR employed repetition of the researcher’s normal recast to move the flow forward. Thereby, CNT4.17.SR may have been unaware of using past tense. NR3.74.SR may have on the other hand been conscious of what the researcher focused on at that time, when repeating the recast to facilitate his/her speech continuation.
6.4.3 Limited Planning

6.4.3.1 Knowing What to Say Just before Speaking

The time pressure and ongoing flow of the speaking narrative task may have added difficulties for students. This may have been why some students raised the issue of little room for planning. Although pre-task Chinese summary reading was given to students before each session, some students recalled knowing what to say just before the moment of speaking, speaking what they could think of at the moment of speaking, interpreting meaning at the same time of speaking, and still thinking at the moment of speaking.

First of all, both CNT1.11.SR (table 131) and CNT1.13.SR (table 132) in appendix XIII displayed stammers and recounted that they were still dwelling on preparing for their speech seconds before they spoke. CNT1.11.SR stammered and hesitated, when s/he was narrating the meaning and pronouncing a word. These obstacles may have been caused by his/her readiness to speak not long prior to the actual speech. Similarly, CNT1.13.SR stammered, repeated words, and demonstrated inarticulate pronunciation of words. These faults may have been because s/he channeled most cognitive efforts to decide what to say in the following immediate second. Furthermore, both CNT1.11.SR and CNT1.13.SR may have therefore missed using past tense in their narratives.

6.4.3.2 Impromptu Speaking

Instead of managing to plan seconds before speech time, some other students only survived their narratives by planning and speaking concurrently. CNT4.17.SR’s (table 133 in appendix XIII) meaning delivery fell short of clarity, when not precisely mentioning that it was Archer God whom the people loved. Listeners may have easily
misunderstood the pronoun “he” after “the people love” as referring to God of the East, because they were close to each other. CNT4.17.SR recalled that s/he was speaking while thinking all at the same pace; this may have thereby obscured his/her alertness to making every meaning clear to the listener. His/her inconsistent use of past tense may have also been resulted. CNT1.22.SR (table 134 in appendix XIII) expressly attributed his/her stammering to using his/her available capacity at that time to think and speak simultaneously, to compensate the lack of planning time beforehand. S/he may have had no capacity left for past tense usage. Stammering and thinking and speaking at the same time were also what CR4.85.SR (table 135 in appendix XIII) focused on doing.

6.4.3.3 Interpreting Meaning at the Same Time of Speaking

Regardless of the help of pre-task planning, some students remained using the speaking task time to interpret the meaning of the story. The following students, CNT1.13.SR (table 136), NR3.23.SR (table 137), NR3.74.SR (table 138), and CR4.53.SR (table 139) in appendix XIII, recounted specifically their occupied attention to decoding the meaning at the concurrent time of speaking. Students were nevertheless given some pre-task time to glance through the pictures and understand the story’s main meaning in their L1. The pre-task preparation seemed insufficient to counter the instantaneous nature of speech communication. The simultaneous cognitive load of meaning interpretation may have affected the students’ use of past tense (CNT1.13.SR), message coherence (NR3.23.SR, NR3.74.SR), and awareness of the corrective feedback (CR4.53.SR).
6.4.3.4 Still Thinking at the Moment of Speaking

Some extreme cases were that students were unable to complete their meaning because of not ready to think as well as speak at that time. CNT1.13.SR (table 140), CNT4.17.SR (table 141), CNT1.33.SR (table 142), CNT2.39.SR (table 143), CR3.57.SR (table 145), and CNT2.75.SR (table 149) in appendix XIII stopped in the middle of their narratives. They revealed that they could not produce any further concrete thoughts and thereby nothing to deliver at that time. This again showed the challenge of deficient rehearsing time in speaking. NR1.69.SR (tables 146-147 in appendix XIII) was also still thinking during speech time, because s/he was unsure about the meaning s/he was able to utter. His/her meaning ended up being confusing. Even when the meaning was clear, NR3.74.SR (table 148 in appendix XIII) was still thinking at that time because of his/her uncertainty about what s/he had said. CR4.53.SR (table 144 in appendix XIII) explicitly expressed his/her wish that everything would have been better off if the task were with more prior preparation time to think. The ongoing thinking may have overshadowed CNT2.75.SR’s use of past tense and CR4.53.SR’s awareness of the corrective feedback.

6.4.4 Pronunciation Problem

Pronunciation of words can obviously cause problems to both speakers and listeners in speech communication, because it is the only medium of message transmission. Students’ ambiguous pronunciation of words can incur comprehension difficulties. Therefore, past tense forms which sound indistinctively from their base forms in speech, for example regular past, past form with no change of vowel from its base form, and past form in the same spelling as its base form, were excluded from the study’s analyses. Although measures had been taken to control the variable of
pronunciation, some students recalled their apparent tense problems as in fact their wrong or uncertain pronunciation—CR3.21.SR (table 150), NR3.23.SR (tables 152-154), CR4.29.SR (table 155), NR3.47.SR (table 160), CR4.53.SR (tables 161-162), CR3.57.SR (tables 163-164), CR4.58.SR (tables 165-166), NR2.61.SR (table 168), CR4.62.SR (tables 169-170), CR4.65.SR (table 171), NR1.69.SR (tables 172-174), and NR4.87.SR (tables 175-179) in appendix XIII. There was utterly no way for the researcher to predetermine whether the student’s incorrect tense use was due to mispronunciation or grammatical mistake at that time. It was therefore inevitable for the researcher to have mistaken students’ use of non-past tense as unawareness to past tense usage.

Pronunciation was demonstrated not only a barrier to assessing students’ actual tense use in speech communication, but also a hurdle to students’ meaning conveyance in term of accuracy and clarity as shown in their recalls—CNT1.22.SR (table 151), NR3.23.SR (tables 152-154), CNT1.33.SR (table 156), CNT2.39.SR (tables 157-158), CNT2.44.SR (table 159), CR3.57.SR (tables 163-164), NR2.60.SR (table 167), CR4.62.SR (tables 169-170), NR1.69.SR (tables 172-174), and NR4.87.SR (tables 175-179) in appendix XIII. These were probably because pronunciation is the key basis of translating meaning and grammar into speech. Moreover, most challenging is when L2 pronunciation is demanded. Some students may also have immersed in thinking about the pronunciation of the intended meaning and hence missed using past tense (CNT1.22.SR and CNT2.44.SR) and attending to the corrective feedback (NR3.23.SR, CR3.57.SR, CR4.62.SR, and NR1.69.SR).
6.4.5 Fading Memory

6.4.5.1 Thinking What has been Said Wrong

Unlike writing, the production of speech has no permanent print to record messages (Cameron, 2001). In other words, what is said would disappear once it is spoken. This may have posed problems to students’ remembrance and thereby ability to compare his/her original narrative with the corrective feedback. The following students encountered the same problem retaining what they had narrated seconds before in their minds. CR3.57.SR (table 180) and the second excerpt of CR4.58.SR (table 182) in appendix XIII showed that the students were aware of the corrective recasts. However, it was the general mistake-pointing function of the corrective recasts that was being noticed. The exact target of the corrective recasts did not appear to have been attended. CR3.57.SR and CR4.58.SR recalled feeling puzzled about their mistake origins, regardless of the emphases on the target forms in both the question prompts and corrective recasts. Therefore, their surface verbal uptake may have been inattentive ones. Students’ fading memory of their passing spoken messages may have impeded their detailed note of the feedback focuses. The fading memory effect also affected CR4.58.SR’s (table 181 in appendix XIII) ongoing speech, as shown in line 7. S/he remembered having been spending effort on carefully tracing back and monitoring his/her previous utterances while narrating. This may have contributed to his/her broken meaning.

6.4.5.2 Ephemeral

Some other students were aware that the corrective feedbacks given were targeting past tense. However, they had no idea that they had in fact committed mistake concerning past tense usage in their narratives. CR4.65.SR (table 183),
NR3.74.SR (table 184), CR2.80.SR (table 185), and CR4.85.SR (table 186) in appendix XIII recounted that they had the use of past tense in mind at that time; but their actual misuse of non-past tense in speech did not seem to have called their attention at that time, until the intervention of the corrective feedbacks. CR4.65.SR specifically attributed absent awareness of his/her actual use of past tense to his/her fading memory of what s/he had already thought of prior to speaking. This suggested that the fading memory effect in speaking not only overcame students’ remembrance of their passing speech, but also their passing preparation for the forthcoming speech. CR4.85.SR (table 190 in appendix XIII) also pointed out the memory issue, especially in L2; and s/he quickly remembered his/her misuse of tense after the prompting hint and before any corrective recast. NR3.74.SR on the other hand associated unawareness of his/her actual use of past tense in speech to the fleeting duration of speech. NR4.84.SR (tables 187-189 in appendix XIII) even mistook the normal recast as advising him/her to lower the pace to produce more articulate speech. S/he later recalled having forgotten the interlocutor’s use of verb in the normal recast, and may have therefore failed to take up the correction. This may have been due to the short-lived duration of passing speech, by either the student or interlocutor.

6.5 Meaning and Form Competition

With the task demand and speaking modality effects, students’ cognitive capacities may have been strained. Their simultaneous attention to meaning and grammar in communicating the story narrative may have been interfered as a result. Students, as shown in tables 191-195 in appendix XIII, experienced conflicts between their attention to meaning and grammar. This occurred and was reported as an issue, no matter whether they noticed their past tense mistakes, corrective feedbacks, the use
of past tense, and exhibited verbal uptake or not at the time of the feedback sessions. This section analyses the different kinds of competitions between students’ attention to meaning and form in their narratives.

6.5.1 Meaning Overrides Form

The first type of meaning and form attentional competition analysed is students’ recall of their focusing on meaning or vocabulary more than grammar. This issue of meaning attention overriding that of grammar can be further dissected into the following different categories.

6.5.1.1 Focus on Meaning Mainly

Some students recounted that they were solely dwelling on the meaning of the current pictures; and may have thereby exhibited some imperfection in their narratives at that time. For instance, CNT1.11.SR (table 196 in appendix XIII) was engrossed in verbalizing the meaning appeared to him/her on that picture; and this may have consumed most of his/her attention that s/he became prone to his/her Cantonese L1 transfer of legitimate deletion of verbs. CNT2.46.SR (table 197 in appendix XIII) on the other hand was bewildered by the names of the different characters appeared concurrently on the picture, because of his/her shift of attention to the meaning of that picture and words that could be used to communicate it. CR4.58.SR (table 198 in appendix XIII) was even stranded by determining the exact meaning of the present picture; and hence s/he seemed disorganized and still being preoccupied with composing the meaning at that time. Lastly, both CNT2.75.SR (table 199) and CNT4.77.SR (table 200) in appendix XIII were trying to trace back to and consult with the pre-task Chinese summaries, for advising the meaning of the current pictures that they were defeated by at that time. This concentration on recalling the meaning
from the pre-task advice may have interfered their smooth-spoken narratives. No clear
signs however showed whether these students were aware of past tense at that time;
so these stimulated recall excerpts may only suggest their intense focus on meaning.
Subsequent sub-sections analyse other excerpts which suggest students’ concurrent
awareness of past tense or not.

6.5.1.2 Focus on Meaning Mainly, Aware of Past Tense

Some students recollected having been aware of their past tense mistakes, the
use of past tense, or having used past tense in their narratives, while pondering on the
meaning simultaneously. Students, such as NR1.10.SR (table 201), NR3.43.SR
(tables 204-205), CR1.55.SR (table 206), CR3.57.SR (table 207), CR4.62.SR (table
208), CR4.65.SR (table 209), and NR3.74.SR (table 213) in appendix XIII, noticed
the corrective feedbacks on their past tense mistakes, and recalled initial attention to
the use of past tense having been drawn away by attention to the pictures’ meaning. In
particular, NR1.10.SR may still have been contemplating construction of the meaning
after taking up the feedback. CR1.55.SR also stayed captured by the picture meaning
in the subsequent narrative. CR4.62.SR however did not show any verbal uptake and
interpreted his/her subsequent mistake of past tense as mistake of vocabulary choice.
Data about taking past tense mistakes as a vocabulary matter will be further analysed
in the last subsection. Some other students, such as CNT4.17.SR (table 202),
CR2.37.SR (table 203), NR1.69.SR (table 210), CNT1.71.SR (table 211), NR3.74.SR
(table 212), and CR2.80.SR (table 214) in appendix XIII, exhibited the use of past
tense in their narratives, while their primary attention was on the story meaning. Both
NR3.74.SR and CR2.80.SR further revealed that they put meaning in a higher priority
than grammar, while having been able to manage the use of past tense in their narratives.

6.5.1.3 Focus on Meaning Mainly, Unaware of Past Tense

In contrast, some students clearly articulated their concentration on the meaning of the pictures, but no clear recall of having noticed their past tense mistakes, the use of past tense, or sign of using past tense in their narratives. Concerning noticing feedbacks on past tense, NR1.10.SR (table 215), NR2.61.SR (table 221), CR4.85.SR (table 224), and NR4.87.SR (table 225) in appendix XIII did not reminisce having detected the feedbacks or their mistakes. Moreover, all of them viewed meaning focus heavily at that time. The difference is that NR1.10.SR did not take up the target form of the feedback accurately; whereas the other three students displayed correct uptakes. Their correct uptakes could have been mechanical repetitions or not. For students who did not receive any feedback on past tense, CNT1.13.SR (table 216), CNT1.22.SR (tables 217-218), CNT2.40.SR (table 219), CNT2.46.SR (table 220), and CNT4.77.SR (tables 222-223) in appendix XIII jointly showed no clear use of past tense in their content narratives, and predominantly attended to meaning at that time.

6.5.1.4 Focus on Vocabulary Mainly

Similar to the first subsection, some students showed no overt evidence of being aware of the corrective feedbacks on their past tense mistakes, the use of past tense, or actually using past tense in their narratives. Students' focuses analysed here were however gathered on the tool of conveying meaning—use of vocabulary or content words (VanPatten, 1996, 2004).
Having problems retrieving L2 words during their narratives had been recalled by NR1.10.SR (table 226), NR1.69.SR (table 236), NR4.84.SR (table 238), and NR4.87.SR (table 239) in appendix XIII. This may have resulted in NR1.10.SR’s stammering and struggling before finally producing the word “teach”; NR1.69.SR and NR4.84.SR’s repeating the same words and meaning; and NR4.87.SR’s use of fillers showing hesitation and struggle with meaning delivery, leading to his/her broken narrative. Some other students, such as CNT1.11.SR (table 227) and NR2.60.SR (table 233) in appendix XIII, encountered difficulties with proceeding with the meaning because of finding words for the progress of the narrative. Their narratives may have thereby contained repetition of words and syllable pronunciation. CNT1.13.SR (table 228), NR3.43.SR (table 230), and CR1.55.SR (table 231) in appendix XIII on the other hand recalled being indecisive with which word to use at that time. This may have led to the former two producing incomplete meaning, and the latter one with broken narrative. Being not confident enough with the words in mind or used was told by CR4.29.SR (table 229), CR4.58.SR (table 232), and CNT1.71.SR (table 237) in appendix XIII. They may have then exhibited incomplete meaning and broken narratives. Even when students may have had the meaning in mind, CR4.65.SR (table 234) and NR1.69.SR (table 235) in appendix XIII revealed that they were defeated by the means of translating the meaning. Both of them conveyed no concrete meaning at that time. CR4.65.SR seemed frustrated and stuck with words already said; and NR1.69.SR could only contribute minimal production through filler or linking word “and then".
6.5.1.5 Focus on Vocabulary Mainly, Aware of Past Tense

A few students in the stimulated recall subset showed signs of awareness of the corrective feedbacks on past tense, use of past tense, and actually using past tense in their narratives, while largely attending to the use of vocabulary for their narratives. CNT4.17.SR (table 240), CR3.21.SR (table 241), CNT1.33.SR (table 242), NR1.69.SR (table 245), and CNT2.75.SR (table 246) in appendix XIII commonly applied past tense to their narratives, and bore the proper use of vocabulary to fulfil the narrative meaning in mind. However, they did not explicitly recall their awareness of past tense usage. NR2.34.SR (table 243) and NR2.61.SR (table 244) in appendix XIII contrastively articulated their awareness of their past tense mistakes besides their verbal uptake of the normal recasts. Their attention to grammar did not appear to have been attenuated by their concurrent efforts of searching for the appropriate words.

6.5.1.6 Focus on Vocabulary Mainly, Unaware of Past Tense

Although NR2.34.SR (table 250), NR3.47.SR (table 255 last excerpt), CR4.58.SR (table 256), NR2.61.SR (table 257), and NR1.69.SR (table 258) in appendix XIII indicated their verbal uptakes of the corrective feedbacks, no recall of noticing their mistakes can be used to insightfully unveil their past tense awareness. The uptakes could have been pure mimicking or not, especially when split attention to finding a word for the meaning at that time stood in the way. NR3.47.SR (table 255) in appendix XIII also exhibited other focuses in his/her first three excerpts. S/he did not seem to have taken up the previous three different normal recasts; this may have been affected by his/her meaning focus, grammar focus, and vocabulary focus respectively. Without the intervention of feedback on past tense, students’ awareness of past tense usage can only be based on their actual use of past tense in their
narratives. However, CNT1.11.SR (tables 247-248), CNT1.13.SR (table 249), CNT2.39.SR (table 251), CNT2.44.SR (table 252), CNT2.46.SR (tables 253-254), and CNT4.77.SR (table 259) in appendix XIII did not show any clear use of past tense. The omission of past tense, as well as their unfinished narratives, may have been caused by their attention fixated on the use of vocabulary at that time.

6.5.1.7 Focus on Polishing Vocabulary Mainly

Students’ stimulated recalls in this subsection also demonstrated their main involvement with thinking the use of vocabulary, as well as no obvious indication of their use of past tense or not in their narratives. However, the difference is that these students attempted a higher-level of thinking. CNT1.11.SR (table 260), NR3.23.SR (table 261), and CNT2.44.SR (table 262) in appendix XIII recounted their dissatisfaction with the words they used at that time and expressed wishes to refine them. They were somehow unable to produce their better versions of word use. This may have resulted in their stammering and broken narratives.

6.5.1.8 Focus on Polishing Vocabulary Mainly, Unaware of Past Tense

There was on the other hand a case which unambiguously showed the student’s use of non-past tense in his/her own narrative, and even in his/her verbal uptake of the corrective feedback. CR1.55.SR (table 263) in appendix XIII displayed struggle with using the noun phrase “all the animals” or “most of the animals” in his/her original narrative. This may have consumed his/her attention on using past tense for “feel”. His/her inaccurate uptake of the corrective recast, remained using “feel” there, and recall of regretting his/her final choice of using “most of the animals” substantially suggested his/her unawareness of the feedback and the use of
past tense. Self-monitoring the appropriateness of his/her own choice of words at that time may have been the reason.

6.5.1.9 Tense Mistake as Vocabulary Mistake

In addition to taking up the target forms of the corrective feedbacks given and placing concentration on the use of vocabulary, some students perceived their past tense mistakes as vocabulary mistakes. NR3.23.SR (table 264), CR4.29.SR (tables 265-266), CR4.58.SR (table 267), and CR4.65.SR (table 268) in appendix XIII illustrated clear verbal uptakes of the feedbacks, but recollected that those were matters concerning correct vocabulary usage instead of tense. In particular, CR4.29.SR and CR4.58.SR further recalled that they were swamped with locating suitable words for the meanings and stressed the use of vocabulary in their other narrative turns. Their concern for vocabulary may have infected their perception of the feedbacks on their past tense mistakes. CR4.85.SR’s case (table 269 in appendix XIII) was quite a unique one, which showed the student’s uptake of the target form but s/he did not seem to have considered the feedback as pointing out his/her mistake. CR4.85.SR regarded the difference between his/her use of form and the target form of the feedback as discrepancy in personal choice of words. There may not have been any awareness of tense suggested.

6.5.2 Meaning-bearing Form

Apart from placing importance on the meaning and vocabulary which conveys meaning in their narratives, some students also diverted their attention to those grammar forms which bear high communicative meaning value. Among the following students who recalled paying attention to forms other than past tense which carry more communicative value than past tense, some revealed their awareness of the
change of tense in the feedback but interpreted it in the wrong sense; and some focused on other grammar aspects.

CR3.57.SR (table 272) and NR4.87.SR (table 277) in appendix XIII noticed there was a difference in tense use between their original output and the feedback. CR3.57.SR showed clear uptake of the corrective recast and detected the feedback was targeting past tense at that time. However, s/he misinterpreted the main verb “have” instead of the auxiliary verb “do” which needed to be changed to past tense. S/he may have been drawn by the higher meaningful value of the main verb “have” than its auxiliary company “do” and thereby made the wrong judgment. NR4.87.SR was also able to take up the feedback and discern the contrast between his/her use of tense and that of the feedback. However, s/he emphasized more on conveying the character Archer God’s emotion at that time; and s/he therefore may have become less meticulous in deciding the use of tense in delivering this meaning or thought the use of tense as redundant in translating the meaning.

NR3.43.SR (table 270), CNT2.46.SR (table 271), NR1.69.SR (tables 273-274), CNT4.77.SR (table 275), and NR4.87.SR (tables 276 & 278) in appendix XIII on the other hand reported in retrospect that they noted their use of other grammar aspects than past tense at that time. NR3.43.SR, NR1.69.SR, and NR4.87.SR showed their awareness of the use of past tense via uptake and actual use. However, they only recalled their focuses on gender reference, passive voice, and preposition and singular verb respectively. Gender reference unambiguously signals the difference between male and female through the use of pronouns “he” and “she”; and passive voice conveys the prominent difference of whether someone’s action is in a giver mode or taker mode through the positioning of the subject and object in a sentence.
NR4.87.SR’s focuses on preposition and singular verb at that time were made
contingent on his/her concurrent use of the object and subject respectively. Their meaningful values may have thereby been higher than tense use which merely depends on conjugation of the verb itself. The higher meaningful values of these other grammar aspects than past tense may have attracted more of students’ attention, regardless of their uptakes and actual use of past tense. For CNT2.46.SR and CNT4.77.SR, they did not receive any feedbacks on grammar and did not use or show any clear use of past tense in their narratives. Both of them recollected that they were pondering on the use of adjective and irregular plural noun at that time, which are of higher meaningful value than tense. Using the wrong adjective or plural noun may easily convey the wrong meaning; but using the wrong tense may still communicate the action across and the time reference can be compensated by time adverbials.

6.5.3 Communicative Strategy

6.5.3.1 Use of Simpler Form to Quickly Communicate

Some other students also recalled attending to meaning mainly, but strategically to ease effective communication at that time. CNT1.33.SR (table 279), CR4.53.SR (table 280), CR4.58.SR (table 281), NR1.69.SR (table 282), and CR4.85.SR (table 283) in appendix XIII all showed clear use of past tense and uptakes of past tense from corrective feedbacks. Moreover, all of them encountered the same experience of being unable to present the meaning in the way they initially wanted and ended up with second best means to manage the narrative meaning. Resorting to a simpler or more accessible way to resolve the obstacles to their communication may have allowed them to weather the difficulties and spare attention to the use of past tense at that time.
6.5.3.2 Knowing Listener will Understand Anyway

Another way of strategically focusing on meaning was shown in CNT4.77.SR’s (table 284 in appendix XIII) stimulated recall, which was attempting any random words in English that came to his/her mind at that time to efficiently facilitate his/her smooth narrative. Different from the students above, CNT4.77.SR did not recall bearing a purposeful meaning or word in mind before approaching an easier narrative. S/he only showed his/her goal of proceeding with the speaking flow, and appeared quite confident that the meaning outcome would be successfully comprehended by the listener. His/her carefree manner may have also spread to his/her use of past tense at that time.

6.5.4 Form Overrides Meaning

Of the exact opposite to the issue of meaning focus overriding form focus, some students manifested their heavier focus on form than meaning. CR2.37.SR (table 285) and CR2.38.SR (table 286) in appendix XIII illustrated clear use of past tense in their narratives. However, they had problems delivering meaning ideas in smooth transition, because of their occupied attention to using past tense at that time. NR3.43.SR (table 287), CR1.55.SR (tables 289-290), and CR4.58.SR (table 291) in appendix XIII exhibited clear use of past tense and uptakes of the past tense feedbacks. Their experience was being unable to move their narratives forward, due to their overwhelming effort to dealing with past tense at that moment over that of meaning. NR3.47.SR (table 288 in appendix XIII) however knew how to handle the accurate meaning at that time but gave up doing that, because of sacrificing his/her effort to take up the past tense mistake.
6.5.5 Simultaneous Meaning and Form Focuses

6.5.5.1 Thinking of both Tense and Meaning

Instead of focusing more on either the meaning or form, some students unveiled their simultaneous engrossment in both meaning and form. CNT2.40.SR (table 292), CR4.58.SR (table 293), NR1.69.SR (table 294), CNT1.71.SR (table 295), and NR4.84.SR (table 296) in appendix XIII commonly exhibited their clear use of past tense in narratives, as well as produced substantial meaning presentation, though the delivery may not have been perfectly fluid. Their comprehensive care for both meaning and form in their narratives may have been fostered by their unbiased or extensive attention to the two aspects at that time.

6.5.5.2 Thinking of Tense when Meaning is Obvious

Other students also noted both meaning and form at the time of narrative, but they disclosed that their simultaneous attention was rather conditional. NR3.43.SR (table 297), NR3.47.SR (table 298), CR4.53.SR (table 299), and CR4.85.SR (table 300) in appendix XIII illustrated overtly their use of past tense in their narratives, and recalled the easy handling of meaning at that same time. Meaning and form were concurrently manageable by them when the pictures conveyed the same meaning, clear vocabulary hint appeared on the picture, easily identifiable features were provided on the picture, and uncomplicated meaning came with the picture respectively.
6.6 Language Development

6.6.1 Interlanguage

6.6.1.1 Tenseless L1 Versus Tense-bearing L2

Students' use of past tense or not in their narratives depends not only on external factors like task demand and task modality, but also on internal factors like students' attention capacity and language development status quo. This section analyses students' recall of spontaneous use of tense in their narratives according to their own language condition at that time. Most students reported here did not deliberately control their use of tense; they manifested in retrospect that they instead allowed the use of whatever grammar or tense came to their minds or they were used to. A few others clearly showed their spontaneous use of past tense, while recalling not having controlled their use of tense and speaking about their key concern for meaning in hindsight.

NR1.10.SR (table 301), CR2.37.SR (tables 302-303), NR3.43.SR (table 304), and NR3.74.SR (table 305) in appendix XIII successfully took up their corrective feedbacks on past tense. In addition to noticing their misuse of tense from the feedbacks, they also confided in their stimulated recalls that they in fact did not pay much heed to the use or need of tense in the narratives. Their unguarded use of tense for the narratives could most likely have been driven by the students' tenseless L1—Cantonese. Students' recalls of initial unawareness of the need of tense use suggested that the feedbacks may have served as awakening calls for them. As shown in appendix XIII, CNT4.17.SR (tables 306-307), CR2.38.SR (table 309), CR3.57.SR (table 310), NR1.69.SR (table 311), and CNT2.75.SR's (tables 308 & 312) mixed use of past tense and non-past tense was also recalled to be caused by their heedless use
of tense. Their spontaneously inconsistent use of tense may have been derived from their tenseless L1 and developing tensed L2 at the same time.

Some students, such as CNT2.44.SR (table 313), NR3.74.SR (table 314), and CNT2.75.SR (table 315) in appendix XIII, however showed clear automatic use of past tense in their narratives. Notwithstanding their use of past tense, they recalled the same circumstance of their heedless use of tense at that time. CR4.65.SR (table 316), NR1.69.SR (table 317), and CNT1.71.SR (table 318) in appendix XIII on the other hand recollected their focuses on meaning and no report of focuses on tense, though they showed clear automatic use of past tense. All their use of past tense may have been occasions of their developing tensed L2. Lastly, CNT1.13.SR (tables 319-320 in appendix XIII) did not show any sign of using past tense. Moreover, s/he exhibited use of verbless narrative, which was obviously a L1 transfer from Cantonese. The double influence from his/her L1, tenseless and verbless, may be further affirmed by CNT1.13.SR’s recall of mindless narrative at that time.

6.6.1.2 Verbless L1

This section further illustrates students’ narrative occasions of verbless L1 transfer, and their recollections of being unaware of it as well as aware of it. Students who were unaware of their L1 influence on their actual L2 narratives recounted that their focuses were projected on other aspects at that time. CNT1.11.SR (table 321), CNT4.17.SR (table 323), CNT2.44.SR (table 325), CNT2.46.SR (table 326), and CNT4.77.SR (tables 327-328) in appendix XIII placed their attention on the meaning of the narratives. CNT1.13.SR (table 322 in appendix XIII) on the other hand cared about the sustainability of the ongoing speaking flow of his/her narrative.
CNT1.33.SR (table 324 in appendix XIII) was exceptional here in the sense that s/he was aware of the L1 transfer impact on his/her verbless narrative.

6.6.2 Unproceduralized L2

6.6.2.1 Naturally Turning Back to Non-past Tense

As illustrated above, students' developing L2 in the study was embodied in their L1 transfer of tenseless use of verbs and verbless sentences. Apart from L1 transfer, students' unproceduralized or out-of-practice use of past tense also exemplified that their L2 was still under development. The following stimulated recall analyses uncovered some students' rusty skill of actually using past tense in their spoken narratives, even though they had learned the use of past tense from school.

Among these students, NR1.10.SR (table 329), CNT2.40.SR (table 330), NR3.47.SR (tables 331-332), CR4.53.SR (tables 333-334), and CR4.62.SR (table 339) in appendix XIII explicitly confided that they had the use and need of past tense in mind at that time, but they ended up using non-past tense. Concerning why or on what condition these students knew the use of past tense but did not take action accordingly, NR1.10.SR, NR3.47.SR, CR4.53.SR, and CR4.58.SR (table 335 in appendix XIII) had no idea about the reason or condition behind. On the other hand, CNT2.40.SR particularly revealed that s/he was focusing mainly on the meaning during the narrative, and that s/he was inclined to using L1 grammar in L2. Apart from focusing mainly on the meaning, CR4.62.SR (table 340 in appendix XIII) also recalled actual online speaking as a condition when s/he missed the use of past tense. Actual online speaking being an intervening condition was also reported by NR2.60.SR (table 338) and CR2.80 (table 341) in appendix XIII. Most of them showed surface uptakes and use of past tense during the feedback sessions,
notwithstanding their recalls of the mismatch between past tense in their minds and narratives. However, the no-uptake occasions of NR3.47.SR (tables 331-332) and CR4.62.SR (table 339), and NR2.60.SR’s (table 338) wrong uptake were exceptions.

6.6.2.2 Inconsistent Use of Past Tense

The above students exhibited their unproceduralized use of L2 past tense by recalling that they could not actualize their established knowledge of past tense in their spoken narratives. The following analyses suggested students’ unproceduralized L2 through another manner—their inconsistent application of past tense.

Students’ inconsistent use of past tense can be rendered by their recollection of using past tense in previous narrative occasions but not in the current one. NR1.10.SR’s stimulated recalls not only advised his/her inability to actualize the use of past tense at that time as shown in table 329 in appendix XIII, but also his/her inability to keeping the consistent use of past tense when s/he could actualize its use. In table 342 in appendix XIII, s/he recounted that was the time when s/he missed using past tense regardless of previously remembering using it after the preceding normal recast. CR3.21.SR’s (table 343 in appendix XIII) use of past tense in some occasions but not others was illustrated in his/her narratives as well as stimulated recall. S/he kept using past tense until his/her mistake, which was accompanied by zero uptake of the corrective recast. His/her mindlessness of maintaining the same tense pattern was further indicated in his/her recall. S/he also spelled the reason of his/her out-of-practice use of past tense. CNT1.33.SR (table 344 in appendix XIII) also reminisced his/her unstable use of past tense. His/her use of “had” could have been “have”, which was difficult to be determined given their same vowel. If that were the case, the excerpt’s corresponding recall would have echoed precisely the

345
student’s variable use of past tense. Some other students’ narratives also explicitly showed their variable use of past tense. CR4.53.SR (table 345), CR3.57.SR (tables 346-347), CR4.65.SR (table 348), and NR3.74.SR (table 349) in appendix XIII commonly showed a combination of initiating the use of past tense plus non-past tense in their narratives. Moreover, they all acknowledged their careless mistakes of missing some occasions of using past tense while having succeeded in others.

6.6.2.3 Grammar-conscious

Chapter 2 has discussed the difference between declarative and procedural knowledge in the language skill approach. Being conscious of grammar or the use of past tense in the present study while using or not using it in production may reveal predominantly students’ L2 declarative knowledge, but unproceduralized or not yet automatized L2 in other words. On the contrary, students’ L2 procedural knowledge may be translated through their automatized use of past tense with their unawareness of using it while speaking. The following students’ stimulated recalls showed three different ways of suggesting their unproceduralized use of past tense at that time.

Students like CNT1.11.SR (table 350), CR3.21.SR (table 351), CR4.29.SR (table 359), CNT1.33.SR (table 362), CR2.38.SR (table 363), NR3.43.SR (table 365), NR3.47.SR (table 366), CR3.57.SR (table 370), CR4.58.SR (table 371), CR4.62.SR (table 375), CNT1.71.SR (table 378), NR3.74.SR (table 379), NR3.74.SR (table 380), and CNT4.77.SR (table 381) in appendix XIII verbalized their wariness over the use of past tense at the time of their narratives. All of them shared the common point of clearly using past tense at that time, except CNT1.11.SR’s unclear use of past tense, NR3.43.SR’s clear uptake of feedback in addition, NR3.47.SR’s no uptake of the feedback, and CNT4.77.SR’s no clear use of past tense.
NR3.23.SR (tables 352-356), CR4.29.SR (tables 357-361), NR3.43.SR (table 364), CR1.55.SR (tables 367-368), CR3.57.SR (table 369), NR2.61.SR (tables 372-373), CR4.62.SR (table 374), CR4.65.SR (table 376), and NR4.84.SR (table 382) in appendix XIII on the other hand showed awareness of their tense problems from their corrective feedbacks. All of them exhibited clear verbal uptake of the feedbacks or their mistakes during the session, except NR3.23.SR’s (table 353) wrong uptake, and CR1.55.SR (table 367) and CR4.62.SR’s (table 374) no uptake occasions.

Lastly, CR4.29.SR (table 361) and CR4.65.SR (table 377) in appendix XIII illustrated their mindfulness of past tense at that time by casting doubt over the specific past forms. The former showed unclear use of past tense, while the latter exhibited clear verbal uptake of the corrective recast. All in all, overt production of past tense or tense correction may or may not have suggested students’ awareness.

6.6.2.4 Mistake-conscious

Similar to the above sub-section, students’ concern for their use or mistake of past tense during their narratives was considered representing their yet-to-be automatized use of L2. The difference between these two sub-sections is that the former focuses on illustrating students’ care for the language aspect of grammar in general, while the latter confines the scope more to their alert for their own mistakes. Students’ alert for their own tense mistakes at that time can be substantiated by their stimulated recalls in the following ways.

Feeling uneasy or suspicious of their committing tense mistakes, guarding against their committing mistakes, or automatically aware of their mistakes without any relevant feedback interventions was recalled by CNT1.13.SR (table 383), NR3.23.SR (table 384), NR2.34.SR (table 385), CNT2.46.SR (table 387), CR1.55.SR
(tables 389-390), CR3.57.SR (tables 392, 393 & 395), CR4.58.SR (table 398), NR2.60.SR (table 399), CR4.65.SR (table 403), CNT1.71.SR (tables 404-406), NR3.74.SR (table 407), CNT4.77.SR (table 408), CR2.80.SR (table 410), and CR4.85.SR (table 412) in appendix XIII. Not all of them however exhibited use of past tense in their original narratives. CNT1.13.SR, CNT2.46.SR, and CNT1.71.SR (table 404) did not use past tense; NR3.23.SR, CR4.58.SR (table 398), and CNT4.77.SR did not show any clear use of past tense; and CR1.55.SR (tables 389-390) did not offer any signs of tense-bearing verbs. These showed that no surface output of past tense may not have been an absolute negation of students' awareness of it.

Some students, such as NR3.43.SR (table 386), CR4.58.SR (table 396), and CR4.65.SR (table 402) in appendix XIII, were also aware of their mistakes without any relevant feedback interventions. They were students' mistakes of other aspects, for example use of vocabulary and narrative delivery, although they clearly took up the feedbacks on past tense. Other students became aware of their mistakes of past tense from the tense feedbacks. Among them, CR1.55.SR (table 388), CR3.57.SR (tables 391 & 394), CR4.58.SR (table 397), CR4.65.SR (table 401), and NR4.87.SR (table 413) in appendix XIII verbalized their uptake of the correct past tense usage; whereas CR4.62.SR (table 400), CR2.80.SR (table 409), and NR4.84.SR (table 411) in appendix XIII did not output any corresponding uptake. Verbal uptake may not have been a guaranteed mirror to students' awareness of past tense.

6.6.2.5 Fluent to Use Some Forms Over Others

Apart from performing unproceduralized L2, a few students in the stimulated recall subset showed their selective unproceduralized L2 performance during their
narratives. Some students recounted that they were able to use past tense or not depending on the particular verbs they encountered at that time. For instance, both CR2.37.SR (table 414) and CR2.80.SR (table 417) in appendix XIII pointed out that their ease with certain verbs drove them to use certain tense at that time. CR2.37.SR particularly voiced his/her difficulty with past tense because of its conjugation complication during online production. S/he however revealed such difficulty was manageable in some cases but not others, relying upon how well s/he dealt with the conjugation complication at that time. CR2.80.SR also counted on the verb cases s/he experienced at that time in using past tense or not. S/he however attributed his/her variable use of past tense to the difficulty and familiarity levels of different verbs. CR4.65.SR (table 415) and NR1.69.SR (table 416) in appendix XIII pointedly articulated their feeling comfortable with using certain tense for the verbs they used at that time. CR4.65.SR strategically used past tense by staying in his/her comfort zone of reusing “didn’t”, which s/he remembered using correctly previously; and NR1.69.SR attributed his/her using non-past tense for “see” to his/her intuitive comfort with it. These students showed their selective use of past tense originated from their strength with certain verbs or tense.

6.7 Different Effects of Corrective Feedbacks and Uptakes

6.7.1 Extensive Effect of Feedbacks

Apart from revealing the different external (task demand and speaking modality) and internal (students’ attention competition and language development) factors, students’ stimulated recalls also reflected the different effects of normal and corrective recasts given in their feedback sessions on past tense. This subsection first of all demonstrates the pervasive effectiveness of the recasts. The following students
recalled that they noticed the feedbacks and use of past tense at that time from the corrective effect of the previous feedback turns. In other words, the effect of recasts extended from previous turns to the current turns. CR2.37.SR (table 418) and NR3.74.SR (tables 422-423) in appendix XIII were able to take up the feedbacks; and they recounted that the previous feedback turns served as reinforcement or indicators to help enhance their awareness of the feedbacks at that time. Other students, CR4.58.SR (table 419), CR4.62.SR (table 420), CR4.65.SR (table 421), and NR4.84.SR (table 425) in appendix XIII, on the other hand automatically used past tense or noticed the use of past tense. They also verbalized the lasting effect of the feedbacks they had been receiving at that time. CNT2.75.SR (table 424 in appendix XIII) stood however as an exception, who did not receive any recasts on past tense but his/her prior habitual use of past tense made him/her stay with the tense at that time.

6.7.2 Consistent Effect of Feedbacks

6.7.2.1 Continuous Mistakes

Another kind of effect of recasts that students’ stimulated recalls suggested was their consistent effect. The following students generally expressed their noticing the feedbacks, their mistakes, and the use of past tense from the regular recurrence of the feedbacks and their mistakes. No matter whether the episodes showed that the student did not take up the corresponding feedback (NR1.10.SR—table 426 in appendix XIII), the students took up the corresponding feedbacks (NR1.10.SR—table 426, NR2.34.SR—tables 427-430, CR2.37.SR—table 431, CR4.53.SR—table 432, CR3.57.SR—table 433, NR2.60.SR—tables 434-435, NR1.69.SR—tables 436-437, CR2.80.SR—tables 439-440, and NR4.84.SR—tables 441-442 in appendix XIII), the
students automatically used past tense (NR2.34.SR—table 427, CR2.37.SR—table 431, CR4.53.SR—table 432, NR1.69.SR—table 436, and CR4.85.SR—tables 443-444 in appendix XIII), or the student took up the feedback but did not reformulate the mistake (NR3.74.SR—table 438 in appendix XIII), they all manifested their constant realization of the focus of past tense brought by the reoccurring mistakes and feedbacks.

6.7.2.2 Reminding the Use of Past Tense

The consistent effect of recasts can also be viewed from students’ recalls of being reminded of their mistakes or the use of past tense by the feedbacks, which they were not aware before the cues. As discussed in the section on students’ unproceduralized L2, students’ mistakes of past tense could most likely have been their performance control problems (procedural knowledge). The consistency of the feedbacks may have then helped direct students’ wrong use of tense back to the right track by persistently giving refreshing signals. Regardless of the zero substantial surface uptake at one occasion, NR3.47.SR (tables 449-450), CR1.55.SR (tables 452-453), and CR4.62.SR (tables 457-458) in appendix XIII showed their subsequent use of past tense or clear uptakes of the feedbacks as well as acknowledging the reviving function of the corresponding feedbacks. Others (NR1.10.SR—table 445, CR2.37.SR—table 446, CR2.38.SR—table 447, NR3.43.SR—table 448, CR4.53.SR—table 451, CR4.58.SR—table 454, NR2.60.SR—table 455, and NR2.61.SR—table 456 in appendix XIII) also displayed clear uptakes and use of past tense. Their awareness of the feedbacks and past tense, reflected by the uptakes and their use of past tense, had been further affirmed by their recollection of having been activated by the feedbacks. The consistent patterns of the feedbacks may have shaped
the feedbacks as easy and quick reminders along the way, even though students may not have noticed their mistakes earlier, rather than nerve-racking ones if feedbacks were incidental and unexpected.

6.7.3 Intrusive Effect of Feedbacks

Besides the positive effects of the recasts, some other students considered them as obstructing their original meanings or continuous flow of speaking. NR3.47.SR (table 459) and CR4.58.SR (table 460) in appendix XIII succeeded in taking up the feedbacks, but they complained in retrospect that the feedbacks and the focus on past tense interrupted their concentration on retaining (NR3.47.SR) and expressing (CR4.58.SR) the meanings already in their minds. The intrusive effect of the feedbacks may run counter to the regular effect of the feedbacks that students above reckoned as familiar and facilitative. For example, CR2.38.SR (table 461 in appendix XIII) explicitly viewed the recurrent feedback at that time as abrupt instead of customary, though s/he showed clear uptake of it. His/her mind may have been captivated by something else at that time, and that one corrective recast event may have therefore appeared startling to him/her.

6.7.4 Confirming Effect of Feedbacks

Some other students neither viewed the recasts as facilitative nor disruptive. They however regarded the corrective feedbacks as confirming their original meanings. NR4.84.SR (table 462 in appendix XIII) did not seem to have noticed the corrective function of the normal recast, from both his/her acoustic uptake without any concrete reformulation and stimulated recall. S/he perceived the normal recast as merely the researcher’s attempt to check his/her own comprehension of the student’s exact narrative. The subsequent uptake without reformulation may have thereby been
the student’s gesture of affirming the researcher’s correct reception of his/her meaning. CR4.85.SR (table 463 in appendix XIII) also decoded the researcher’s feedback as not given to his/her use of tense, though s/he took up the feedback. The uptake may have hence been a superficial one. From his/her first response to the feedback procedure, s/he may have been troubled by the preciseness of his/her original meaning, and therefore preoccupied with experimenting the meaning after the initial question prompt. From his/her recall, CR4.85.SR may have then felt relieved after receiving the corrective recast, which s/he identified as acknowledging the viability of his/her original meaning.

6.7.5 Uptake as Practice to Proceduralize

The different effects of recasts on students had been shown through investigating their personal experience with perceiving them at the time of the feedback sessions. Uptake, which was expected to suggest students’ awareness of the recast, had also been regarded by some students as facilitative for them to better interpret the current feedbacks as well as avoid future mistakes. The following students, NR2.34.SR (table 464), CR1.55.SR (table 465), NR2.60.SR (table 466), NR2.61.SR (table 467), and CR2.80.SR (tables 468-469) in appendix XIII, commonly showed fruitful uptakes in the sense that they not only noted the recasts, but also utilized their uptakes to practice or familiarize themselves with the corrections. NR2.60.SR and CR2.80.SR further recalled their treating the practice as precaution against repetitive mistakes.

6.7.6 Uptake as Confirmation

Another function of uptake for students in addition to verbalizing their feedback awareness was recalled to be confirming meaning or their correct use of
tense. CR2.38.SR (table 470 in appendix XIII) took up the corrective recast and revealed that his/her uptake was to assure to the researcher or reinforce his/her own awareness of the correct target form or meaning. NR4.87.SR (table 471 in appendix XIII) also articulated that s/he noticed the different tense s/he used from that of the researcher, and considered his/her uptake as substantiating his/her awareness. These students unveiled the strengthening effect of their uptakes.

6.7.7 Uptake Not Guaranteeing Awareness

As suggested by students’ stimulated recall excerpts throughout this analysis chapter, students’ verbal uptakes may not have been attentive ones since their recalls did not indicate that thoughtfulness was necessarily involved in their uptakes. Same as many such examples, NR1.10.SR (table 472 in appendix XIII) exemplified that the uptake was in fact a quick mimicking for his/her speech to be smooth enough to transit to narrating the originally well-thought meaning in his/her mind. The uptake may have thereby been a peripheral transiting tool; whereas narrating the well-thought meaning was the major task for NR1.10.SR at that time.

6.7.8 No Uptake

The existence of students’ verbal uptakes may be directive in the sense that they may suggest their awareness of the feedbacks. Meanwhile, they may appear ambiguous because they may not necessarily signal students’ awareness. In such case, students’ subsequent extensive use of past tense in other verbs may unquestionably suggest their learning the use of past tense from the previous feedbacks. This more overt way of reflecting students’ uptake or learning of the feedbacks may be a more useful indicator, especially when uptakes are not appropriate during the ongoing flow of speaking or when they are redundant to students.
6.7.8.1 Rule Learning

Students’ extensive use or wide application of the rule of past tense to other verb items has been shown in some prior excerpts analysed in this chapter. Among many of these excerpts, CR2.80.SR (table 473 in appendix XIII) below explicitly voiced his/her awareness of the corrective recast in the stimulated recall, though there was no verbal uptake exhibited at that time. However, his/her subsequent use of past tense for another verb item was shown. The extensive use of past tense may have been CR2.80.SR’s attentive uptake of the previous feedback on past tense.

6.7.8.2 Ongoing Flow

In the case when students noticed that the recasts were targeting past tense but did not verbalize their uptakes at that time, their subsequent use of past tense may have been supportive of their implicit uptakes. Students such as NR3.47.SR (table 474), CR4.62.SR (table 475), NR1.69.SR (table 476), and CR2.80.SR (table 477) in appendix XIII shared the point of no solid uptakes of the feedbacks. Moreover, they commonly recalled that the necessity or urge to continue with the narratives or the speaking flow overrode their initial intention to use past tense or verbalize their uptakes of the feedbacks. The ongoing narrative flow may have been the obstructive force to students’ uptakes, rather than their unawareness.

6.7.8.3 Redundancy

Another occasion where opportunities for students’ immediate verbal uptakes were made futile was students’ consideration that verbal uptakes were unneeded. NR3.47.SR (table 478 in appendix XIII) did not show any response to the normal recast, but s/he recollected his/her well-informed knowledge of the recurrent mistake s/he kept having at that time. Superfluously restating the correction may not have
been essential to NR3.47.SR any more, when s/he had full understanding of his/her chronic mistake. Subsequent use of past tense for other verb items may have been more meaningful to the student.

6.8 Conclusion

The above analyses of the study’s qualitative data show the different themes arisen from and shared by the students’ stimulated recalls of their inner thoughts during the feedback sessions. These qualitative data were not meant to be interpreted quantitatively. However, a rough trend can be concluded from the majority of students within the stimulated recall subset who reported having thought of certain themes during the feedback sessions. Among reported themes such as sequential or continuous picture narrative task demand (1), task familiarity challenge (2), task completion priority (3), task organization priority (4), insufficient time in speaking (5), ongoing flow in speaking (6), limited planning in speaking (7), pronunciation problem in speaking (8), fading memory in speaking (9), meaning attention overrides form attention (10), attention to meaning-bearing form (11), communicative strategy adoption (12), form attention overrides meaning attention (13), simultaneous meaning and form focuses (14), students’ interlanguage (15), students’ unproceduralized L2 (16), extensive effect of feedbacks (17), consistent effect of feedbacks (18), intrusive effect of feedbacks (19), confirming effect of feedbacks (20), uptake as practice to proceduralize (21), uptake as confirmation (22), uptake not guaranteeing awareness (23), and no uptake but extensive use of past tense in other verbs (24), considerable amount of data were related to themes (1), (3), (4), (7), (8), (10), (15), (16), (18), (23), and (24). In particular, themes (10) and (16) were recalled most by the subset. The
qualitative results will be further discussed in the following chapter, together with the study's quantitative results.
Chapter 7 Discussion

7.1 Introduction

This chapter discusses the study’s findings in relation to the research questions and the recurring themes that have arisen in learners’ stimulated recalls. In other words, the two sets of data, quantitative and qualitative, will be brought together. Their convergence will be discussed by confirming results from each data set; and their divergence will be informed by using one data set to interrogate the other. Before the comparing and contrasting, summary of the findings of each data set will be provided. The pattern, correspondences and differences of the qualitative data in relation to the quantitative data will be holistically discussed.

Some explanations will be suggested for the findings of each data set. Firstly, explanations will be provided for the quantitative findings about: the immediate effectiveness of unenhanced recasts in promoting learners’ use of irregular past in their spoken narratives (research question 1); the lasting effectiveness of unenhanced recasts in promoting learners’ use of irregular past in their spoken narratives (research question 2); the immediate effectiveness of enhanced recasts in promoting learners’ use of irregular past in their spoken narratives (research question 3); the lasting effectiveness of enhanced recasts in promoting learners’ use of irregular past in their spoken narratives (research question 4); the comparative effectiveness of unenhanced and enhanced recasts in promoting learners’ use of irregular past in their spoken narratives in the short run (research question 5); and the comparative effectiveness of unenhanced and enhanced recasts in promoting learners’ use of irregular past in their spoken narratives overtime (research question 6). The explanations will be based on past findings discussed in the literature review, learners’ inner thoughts revealed in their stimulated recalls, and the present study’s features.
Secondly, past findings in the literature review and the present study’s features will also be used to explain the thematic findings in learners’ stimulated recalls: task demand, speaking modality, meaning and form competition, language development, and different effects of corrective feedbacks and uptakes. The chapter will conclude by looking at the results more widely with regard to any issues that arise in the field, problems in the data collection and analysis, and other difficulties with methodology.

7.2 Research Question 1

Research question 1 asked whether learners gained from the unenhanced recasts consistently targeting their use of irregular past in their spoken narratives; and if they did, whether the gain was an immediate one. The quantitative analysis in chapter 5 indicated that unenhanced recasts (normal recasts) facilitated learners’ use of irregular past more significantly than content-only feedback (control group) in their immediate narrative post-tests.

This finding provided additional support for previous research suggesting the facilitative effect of normal recasts in fostering learners’ use of past tense in a controlled setting in the short run. For example, similar to Han’s (2002) study, the present study employed consistent supply of recasts, intensive focus on irregular past, individualized attention to each learner’s spoken narrative, and target form already learnt by learners. Moreover, this finding of the present study echoed Panova and Lyster’s (2002) ESL classroom finding that recasts were suitable for learners with low proficiency. It also substantiated Swan’s (2005) suggestion that recasts help expand EFL learners’ limited grammatical foundation for task-based communication due to their minimal exposure to the target language. The Hong Kong EFL learners’ gain from normal recasts, though unenhanced, may have also been due to their habitual
reception of grammar-oriented instruction in their Asian educational culture (Ellis et al., 2001; Ellis, 2009a). About gaining in a short period of time, this finding of normal recasts also seemed to consolidate Ellis’s (2007) claim. He contends that form already present in learners’ knowledge base but exhibiting learners’ deficient mastery of it can be polished by intensive feedback in a short period of time. Unlike enhanced recasts (corrective recasts) in the present study, normal recasts appeared predominantly as positive evidence directly modelling the lexical-based irregular past form. The result answered research question 1, as well as confirmed Egi’s (2007b) finding that the positive evidence in recasts facilitated learners’ learning of lexical item-based form, which only calls for learners’ exemplar memorization.

The result corresponding to research question 1 seemed to have been derived from several interrelated factors besides the mere facilitation of normal recasts. The factors were the controlled operation of recasts, the level-appropriateness of the learners to benefit from recasts, the form-oriented culture of the learners, the prior knowledge of the learners, and the harmony between the nature of the target form and the signature component of normal recasts. All these factors represent the effort of a well-controlled study.

However, some past studies with the control of the setting and the form-oriented EFL background failed to show recasts effective. Iwashita’s (2003) quasi-experimental study showed the short-term benefit of recasts, concurring with the present study’s answer to its research question 1. However, the amount of recast supply was small because of the reticent culture of the Japanese interlocutors. The short-term benefit of recasts resulted may not be representative to allow the conclusion that the controlled quasi-experimental setting may have fostered the effectiveness of recasts. The EFL context of Tsang’s (2004) study of Hong Kong
learners also did not seem to contribute to his learners’ gain from recasts, possibly due to the mixed form and meaning focuses of his classrooms and the lack of training of Hong Kong learners in responding to recasts. Therefore, the overarching factors of the controlled setting and EFL context may not be determinative enough to cause absolute effects to the results of recasts, without considering the specific characteristics of different studies. Nicholas et al.’s (2001) generic claim that controlled settings and form-oriented EFL contexts facilitate the usefulness of recasts as learners’ input may on the other hand be misleading.

The factors of controlled setting and EFL context are significant in the present study to explain the resulted effectiveness of normal recasts in the short run, due to its implementation characteristics. Not only did the present study take the experimental approach instead of an authentic classroom which involves multiple interfering factors (e.g. Lyster, 1998a), but also adopted consistent recasts targeting one form intensively and individually to each learner throughout. Such multi-layered control in the present study may have isolated any possible extraneous influences to the results of recasts, for example the reticent culture or lack of training of Hong Kong learners in responding to recasts (e.g. Iwashita, 2003), the mixed modes of form and meaning-oriented teaching in EFL classrooms (e.g. Tsang, 2004), and the inactive engagement of learners in learning from recasts when given in classrooms with the presence of other learners (e.g. classroom studies). The controlled focus on irregular past could have then easily been retained in learners’ short-term memory immediately after the recast treatment, especially when the form was already learnt by learners. Together with the consistency and intensiveness, the non-intrusive nature of normal recasts in not requiring learners’ responses may have also catered for the learners’ low-intermediate level and novelty with recasts (e.g. Yoshida, 2008). The suitability of the
present study’s specific choice of lexical-based irregular past as the positive evidence in normal recasts may have further favoured the effectiveness of normal recasts in the short run, as well as in the long run.

7.3 Research Question 2

Research question 2 asked whether learners gained from the unenhanced recasts consistently targeting their use of irregular past in their spoken narratives; and if they did, whether the gain was a lasting one. The quantitative analysis in chapter 5 indicated that unenhanced recasts (normal recasts) facilitated learners’ use of irregular past more significantly than content-only feedback (control group) in their delayed narrative post-tests.

This finding of the lasting effect of normal recasts provided reinforcing support for previous research emphasizing the aptness of normal recasts, though implemented in the unenhanced and ordinary way, in facilitating learners’ use of irregular past. The first argument is based on the cognitive aptness of normal recasts to learners. The aptness of the positive evidence element in normal recasts in targeting the lexical or memory-based target form (Egi, 2007b; Yang & Lyster, 2010) of the present study may have compensated the unenhanced saliency of normal recasts. This fit between the feedback and the form may have possibly enhanced learners’ memory retrieval of using irregular past regardless of the passing time. Both Yoshida (2008) and Philp and Tognini (2009) reckon that recasts are tailor-fit for EFL teachers and learners because of their direct and hence time-saving scaffolding for EFL learners, whose cognitive capacity may not allow them to initiate self-correction, during their limited lesson time. The present study’s EFL learners may have been comfortable
with the non-intimidating nature of recasts, and thereby learned well and continued to benefit from recasts over time.

The second argument is based on the pedagogical aptness of normal recasts. Applying DeKeyser's (2003, 2007) skill practice theory, the repeated exposure and responses to the positive evidence given in recasts may help speed up learners’ transfer from declarative to procedural knowledge. Accordingly, the present study’s learners may have had their declarative knowledge of past tense transferred to its procedural or automatized use from consistently receiving direct positive evidence in normal recasts. This may have been why learners could exhibit lasting gain from normal recasts. Moreover, according to Johnson and Jackson's (2006) performance-based training conception in developing skills, the direct positive evidence in normal recasts may have been effective in automatizing learners’ use of past tense because of their immediacy in confronting learners’ online mistakes with the models. Even when learners had not yet automatized their use of past tense regardless of their repeated exposure to the positive evidence in normal recasts, the declarative intonation in normal recasts may have cast a didactic effect on learners (Loewen & Philp, 2006). The didactic effect may have then reactivated their habitual form orientation in class. Their reactivated habitual form orientation in the present experimental study may have helped the lasting effectiveness of normal recasts.

The result corresponding to research question 2 seemed to have been derived from the directness of normal recasts in suiting learners’ cognitive capacity and aiding learners’ automatized use of past tense. The directness of normal recasts in giving feedback to lexical-based irregular past may not only foster the non-intrusiveness of normal recasts to learners’ online performance, but also Chinese learners’ learning because of their rote-memory language learning tradition (Rao, 2002). These may
have favoured the sustainability of the effectiveness of normal recasts across time. Normal recasts may on the other hand be interpreted as being implicit.

Some past studies found that the implicitness in recasts is detrimental to their effectiveness in facilitating learners’ learning. Early L2 classroom studies, such as Lyster and Ranta (1997) and Lyster (1998a), have identified the lower effectiveness of recasts than elicitations in encouraging learners’ self-initiated reformulation. Recasts were found ambiguous as formal feedback because they resembled teachers’ meaning approvals within the interaction flow. The unclear line between the formal and meaning orientations in recasts is largely attributed to their integration with, instead of break from, the ongoing communication. This conclusion is based on the comparisons made between recasts and elicitations (e.g. Lyster & Ranta, 1997; Lyster, 1998a; Panova & Lyster, 2002), and between recasts and prompts (e.g. Ammar & Spada, 2006; Ammar, 2008; Ellis et al., 2006; Lyster, 2004). These studies illustrated the short and long-term superior effect of elicitations or prompts than recasts because of the comparative implicitness of recasts in drawing learners’ attention to or reflection on their errors. However, as discussed in sections 2.3, 2.4 and 2.6 in chapter 2 and section 3.2 in chapter 3, the problem of using uptakes as measurement, the mismatch between the explicitness of the measurement test and that of the feedback being tested, and the frequency variable may have interfered with the validity of the results in these studies.

These interferences have been controlled in the present study to secure the result validity. According to Bardovi-Harlig (2000), the acquisition of the implicit knowledge of tense-aspect-meaning mapping is a gradual and time-consuming process. The reliance on an immediate kind of measurement, as in the use of uptakes in L2 classroom studies, to gauge the effectiveness of recasts on past tense may not be
applicable, because of the time implicit knowledge development may take. Therefore, the present study made use of a delayed post-test to cater for the time that learners’ implicit knowledge development may need. Normal recasts may have thereby been able to exhibit effectiveness in a fair way, regardless of their comparative implicitness against other response-eliciting feedbacks.

Moreover, the use of spoken narratives eliciting and testing learners’ feel for the use of past tense in the present study may have also offered a fair measurement of learners’ learning from normal recasts, because of the implicitness match between the test and normal recasts. This may have resolved the flaw of using explicit knowledge tests to measure the effectiveness of recasts and consequently biasing the effectiveness of prompts in those studies comparing prompts with recasts. Concerning the frequency variable, the present study supplied normal recasts consistently, in similar amount to the supply of corrective recasts as illustrated by the Mann-Whitney U Test in chapter 5. This may have also resolved the flaw of unfairly supplying more frequent prompt-like input to learners in those prompt-recast studies.

The use of the different kinds of feedbacks in equal consistency may have also addressed McDonough’s (2007) concern of minimizing the difference between the two feedbacks before comparing them, to reach a fair and convincing comparison. Apart from equalizing the consistency between normal recasts and corrective recasts, the present study also made them both involve one error change (Loewen & Philp, 2006) and reduced reformulation by not repeating the entire utterances of learners (Sheen, 2006). All these feature controls between the two feedback types may have resolved the biased delivery of prompts over recasts in those prompt-recast studies, and avoided discriminating the implicitness in normal recasts too much from the other feedback type. A fair measurement and delivery of normal recasts and another
feedback type in the present study may have disambiguated the view of learners’ lasting gain from normal recasts. With the appropriate controls in the study, the only difference in saliency can then be sharply singled out for the more focused comparison in the present study.

7.4 Research Question 3

Research question 3 asked whether learners gained from the enhanced recasts consistently targeting their use of irregular past in their spoken narratives; and if they did, whether the gain was an immediate one. The quantitative analysis in chapter 5 indicated that enhanced recasts (corrective recasts) facilitated learners’ use of irregular past more significantly than content-only feedback (control group) in their immediate narrative post-tests.

This finding of the immediate effect of corrective recasts provided support for previous research illustrating the value of perceptual salience in giving recasts. As early as Schmidt’s (1990, 1993) theory, saliency has been hypothesized as crucial to learners’ language learning. In Ellis and Sheen’s (2006) specific review of recasts, saliency in bringing the target form to learners’ immediate attention is stressed as fundamental for recasts to project effectiveness. With regard to learners with low proficiency and developmental unreadiness in their study, Panova and Lyster (2002) consider that the implicitness of negative evidence in recasts may not be salient for them to learn anything. Holding a relative view of developmental readiness, Ellis (2007) predicts the short-term effectiveness of intensive feedback in targeting forms which are already learned by learners but deficient in their mastery. Ellis’s (2007) conception may also explain the immediate effect of corrective recasts on the present study’s learners who had learned past tense but lacked proficiency in using it. This
seems to relate largely to learners’ prior knowledge or developmental readiness, regardless of the saliency of the feedback used.

Back to the factor of saliency, which drew the key difference between normal and corrective recasts, it has been credited as driving learners’ active online processing of the negative evidence in Yang and Lyster’s (2010) evaluation of prompts. Yang and Lyster specify that prompts are more suitable than recasts to target regular past. They explain that the higher saliency of prompts in pushing learners’ self-repair may motivate learners’ active processing of regular past, which has low communicative value and voiceless -ed morphology, leading to heavy computational load on learners’ online speech performance. The preceding repetition-like questions in corrective recasts served the similar function to prompts in eliciting learners’ initial attempt of self-repair, and thereby propelling their active online processing of their errors and the target form. The additional saliency in corrective recasts of having preceding prompts may have expedited learners’ processing of the negative evidence, and resulted in their immediate gain.

Another reference raising the saliency credit of prompts which can be used to explain the immediate effect of corrective recasts is Han and Kim’s (2008) recommendations. They recommend that using prompts to foreground recasts beforehand may motivate learners’ repair and following incorporation of the reformulation. This additional means to make recasts outstanding may have matched with learners’ “natural, perceptual tendency” (Han & Kim, 2008, p.5) towards salient and easy-to-capture input. The present study’s learners may have been spontaneously drawn to learn from corrective recasts and used the form instantaneously after the treatment.
The proper use of salient recasts to target non-salient forms may also explain the present study learners' immediate gain from corrective recasts. Although the different irregular past tokens appear as individual lexical items (Ellis, 1987), they must undergo the morphological transformation process of verb conjugation from their base forms beforehand. According to Mackey et al. (2000), recasts targeting morphosyntactic forms were less perceived as corrective feedback because of their lower communicative value than lexical, semantic and phonological forms. The use of salient recasts, corrective recasts, in the present study may have immediately secured learners' attention to and use of the morphological target form, regardless of the competition with past time adverbials, which carry more communicative value. Furthermore, the combination of negative and positive evidence in recasts is considered by Egi (2007b) as facilitating morphosyntactic hypothesis testing and rule generalization. The inclusion of repetition-like question prompts and subsequent recasts in corrective recasts may have exactly delivered both negative and positive evidence to learners. Learners may have succeeded in their past tense rule generalization after the double reinforcement, as shown in their immediate performance improvement.

The result corresponding to research question 3 may have been derived from the additional help of prompts to recasts. The same result may have been yielded without the subsequent recasts in the implementation of corrective recasts. The present study may turn out to be comparing normal recasts and prompts, rather than normal recasts and corrective recasts. However, as shown in the Mann-Whitney U Test in chapter 5, the delivery of corrective recasts (question prompts plus recasts) was given in similar amount to that of normal recasts. Corrective recasts were only counted when both the question prompts and the subsequent recasts were given. In
other words, if learners had already reformulated their errors after the question prompts, corrective recasts would not have been delivered. This suggested that the immediate gain of the corrective recast group was obtained from the collaborative contribution of question prompts and recasts. Such collaborative effort may have formed the explicitness of corrective recasts and promoted learners’ gain without delay.

Scheffler (2008) appreciates the use of explicit feedback because it may enhance EFL learners’ form analysis. He explains that the incorporation of explicit feedback with traditional instructions and communicative tasks may cater for learners’ non-linear internal syllabus and inability to be self-responsible for target language code analysis in communication respectively. He also reckons that the explicit reminder of declarative knowledge in explicit feedback may facilitate learners’ procedural knowledge development. Scheffler’s view of explicit feedback coincides with N. Ellis’s (2005, 2006) interpretation that saliency may boost the interface between learners’ explicit and implicit learning.

However, Krashen (1985) disagrees and contends that there is no such interface between explicit and implicit knowledge. Accordingly, the saliency in corrective recasts may only drive learners’ explicit learning of reformulating their errors in the feedback instances, but may not facilitate their subsequent implicit learning of automatically using the target form in other instances. Moreover, the explicit consolidation or reminder of the declarative knowledge of past tense through corrective recasts may lead to learners’ reliance on the feedback, and inability to ever become self-responsible for form analysis. The worst scenario is that learners themselves become irritated by the explicitness of the feedback in breaking their meaning expression (Truscott, 1999, 2007).
The present study learners’ relative grammatical difficulty with past tense and their L1 interference may justify the advantage found of the explicitness in corrective recasts. Ellis (2006a) suggests that the same grammar feature comprises its different difficulty levels based on whether it is implicit or explicit knowledge to learners. In the example of Chinese learners, as quoted by Rothman (2007), they often possess sophisticated explicit knowledge of grammar and do not have difficulty with grammar rules. They however have deficient implicit knowledge of using grammar in production, especially when their L1 grammar does not share anything with the L2 grammar in focus. When their entrenched L1 automatizes learners’ use of L1 instead of L2, salient feedback on L2 may be effective in offering positive evidence for the use of L2 and negative evidence for the competing L1 form (MacWhinney, 1987).

Adopting the perspective of the skill learning approach, both Leeman (2007) and Scheffler (2008) argue that increased use of explicit feedback may help inject L2 declarative knowledge into learners’ automatized L1 usage as reminders to destabilize learners’ fossilization (Selinker, 1972). The present study learners may have needed and benefited from the salient feedback as strong input cues to remind their already-learned use of past tense, to destabilize their L1 use of tenseless verbs in their L2 spoken narratives. As shown in their immediate gain, the strong input cues in corrective recasts may have succeeded in turning learners’ automatized L1 into weaker cues in the competition for associating with the same meaning (MacWhinney, 1987).

7.5 Research Question 4

Research question 4 asked whether learners gained from the enhanced recasts consistently targeting their use of irregular past in their spoken narratives; and if they
did, whether the gain was a lasting one. The quantitative analysis in chapter 5 indicated that enhanced recasts (corrective recasts) did not facilitate learners’ use of irregular past more significantly than content-only feedback (control group) in their delayed narrative post-tests.

This finding of the non-lasting effect of corrective recasts did not seem to provide support for previous research advocating the effectiveness of explicit feedback. Studies showing the development of automatized and hence lasting implicit knowledge from exposure to explicit input were also not supported. The facilitation of corrective recasts may have only been momentary stimulation to remind learners explicitly the use of their unproceduralized past tense. Bardovi-Harlig notes that the implicit knowledge of mapping tense-aspect with meaning in production usage takes a gradual and time-consuming process to develop. Learners’ benefit from corrective recasts may have lasted if the feedback were given across time for learners to have an extended exposure. The explicitness in corrective recasts delivered across a rather short treatment length in the present study may have only captured learners’ short-term memory and facilitated their immediate learning of past tense. If the target form were some explicit knowledge of conjugating verbs into their past tense in fill-in-the-blank exercises, which takes a relatively shorter time of development, the effect of the short treatment length of corrective recasts may have lasted.

The treatment length of normal recasts was also as short as that of corrective recasts; normal recasts however exhibited a lasting effect. The difference may have been derived from the intrusiveness of the preceding question prompts in corrective recasts. The demand for learners’ self-initiated attempt of error correction may have conflicted with the present study learners’ cognitive style. Hong Kong learners are accustomed to passively receiving input because of the teacher-fronted and exam-
oriented education culture of Hong Kong. As Spada and Lightbown (2008) argue, whether or not the explicitness in feedback is obvious to learners depends on their perception towards it. Hong Kong learners may not be sensitive to input requiring active participation to correct their own errors, due to their habitual education of passively receiving teachers’ positive evidence. The push by the question prompts may have become awkward to the present study learners; and learners may not have learned effectively from corrective recasts enough to exhibit their lasting gain. The immediate effectiveness of corrective recasts may have only been due to the latest overwhelming effect of feedback consistency and intensiveness.

According to Dekeyser’s (2003, 2007) skill practice approach, the question prompts in corrective recasts may have interrupted learners’ repeated exposure to the positive evidence in recasts and their practice continuity to proceduralize the target form. From the viewpoint of Skehan’s (1998) information-processing model, the demand for learners’ effort-taking self-correction before receiving any direct help from recasts may have also reduced learners’ free cognitive capacity to attend to the subsequent recasts. Proceduralization may have as a result been obstructed. Borrowing Johnson and Jackson’s (2006) performance-based learning, learners’ proceduralization may have also been obstructed by corrective recasts, due to the lack of opportunity to immediately confront learners’ errors with the target form in recasts during online performance. The disruption caused to the proceduralization continuity, which may be a crucial time-taking prerequisite to develop implicit knowledge, may have consequently disabled the lasting effectiveness of corrective recasts.

Corrective recasts may not only have disrupted learners’ proceduralization, but also the harmony between exemplar-based target forms and positive evidence feedbacks. Yang and Lyster (2010) conclude that recasts, with unclear negative
evidence, are suitable to target exemplar-based grammar which bears no clear rule and only requires learners’ retrieval of holistic items. They specifically quote the example of irregular past to illustrate that it is fit to be directly targeted by the positive evidence in recasts, because it does not necessitate as much online computational processing as regular past before learners can use it. The question prompts in corrective recasts signalling negative evidence may have been redundant when targeting irregular past. The negative evidence property of the question prompts may on the other hand be suitable to facilitate learners’ use of regular past. However, the present study did not investigate learners’ use of regular past because it is difficult to disentangle their phonological ambiguity in speech. The present study may have failed to demonstrate the lasting effect of corrective recasts due to their unsuitability to target irregular past. The immediate effectiveness of corrective recasts may have come from other effects such as the latest overwhelming influence of feedback supply consistency.

Pedagogically, as noted by Loewen and Philp (2006), the interrogative intonation of the question prompts in corrective recasts may have given a misleading message to learners that they were given the choice between error correction and meaning confirmation. The present study learners may have interpreted the question prompts as requesting for their meaning confirmation and the following recasts as the researcher’s reciprocal confirmation of learners’ meaning. This reasoning was based on Hong Kong learners’ convention of receiving formal input in a didactic way, and the usual communicative function of interrogation. The immediate post-tests may have shown the deceptive effectiveness of corrective recasts from the practice effect of doing a similar task immediately. The delayed post-tests, with the test of time, may
have on the other hand unveiled learners’ misinterpretation of corrective recasts as merely meaning confirmation.

Another explanation for learners’ disregard of their errors and hence the non-lasting effectiveness of corrective recasts is based on Nassaji’s (2009) study. Nassaji chose to study learners’ errors of new forms instead of learned forms, because of the fear that learners’ errors of learned forms may have only been their slip of the tongue. He reckons that targeting learners’ slip of the tongue may not be effective because learners may not consider the errors as errors and learn from the redundant feedback. The present study’s focus on learners’ performance-based errors of past tense or learned but unproceduralized form of past tense may not have made learners seriously consider their errors as errors. Learners’ immediate gain from corrective recasts may have been derived from the fresh alert for guarding against their slip of the tongue shortly after the feedback session. With the test of time, the delayed post-tests may have unveiled learners’ disregard of the original intent of corrective recasts due to their redundancy in possibly targeting their slip of the tongue. The match between the non-intrusiveness of normal recasts and learners’ passive cognitive style may have however overridden the redundancy of normal recasts in possibly targeting their slip of the tongue at that time.

With the consistent and explicit supply of corrective recasts in the present study, the aforementioned redundancy of recasts targeting learners’ learned form could have been remedied. Both Swan (2005) and Scheffler (2008) contend that increased use of explicit feedback may remind and reinforce EFL learners’ proper use of grammar in communication. However, the increased exposure to explicit feedback may lead to EFL learners’ over-reliance on feedback during online performance. When time passes and learners are left alone with no concurrent feedback or feedback
given not long ago, learners may lose the support that they used to enjoy effortlessly. They would need to spend more of their own cognitive capacity to use past tense when recasts had not been present for quite some time in the delayed post-tests (Trofimovich, Ammar & Gatbonton, 2007). Moreover, Hong Kong learners often lack learning independence due to their teacher-fronted education. The independence to exert their own cognitive capacity to use past tense in the delayed post-tests may have consequently become alien to them.

The above justifications for the non-lasting result of corrective recasts can be concluded as about the clash between the implementation of corrective recasts and the nature of knowledge that learners need to acquire and learners’ cognitive style. The present study however attempted every possible control to avoid biased investigation of the respective effect of normal and corrective recasts. Control measures adopted for this purpose were: easing learners’ attention to recasts by choosing a form that they were developmentally ready for, encouraging their use of past tense with the use of tasks which may easily elicit their use of past tense, using proper measurement tests to gauge their development, and minimizing their cognitive burden during online performance. The aforementioned clash possibly incurred by corrective recasts may have superseded the controlled effect of the present study.

Moreover, the present study also made the implementation of normal and corrective recasts as similar as possible except their saliency, for example both were given consistently and intensively. The difference between corrective and normal recasts could then be kept marginal, for an accurate examination of their crucial difference in saliency. However, from the insignificant comparative results of normal and corrective recasts, these control measures of minimizing their differences may
have turned out to be affecting a discerning view of the difference between the two feedbacks.

7.6 Research Questions 5 & 6

Research questions 5 and 6 asked whether learners gained more from the enhanced recasts consistently targeting their use of irregular past in their spoken narratives than their unenhanced counterpart. If they did, whether the gain was immediate and overtime respectively. The quantitative analysis in chapter 5 indicated that enhanced recasts (corrective recasts) did not facilitate learners’ use of irregular past more significantly than unenhanced recasts (normal recasts) in both their immediate and delayed narrative post-tests.

These findings of the insignificant differences between normal and corrective recasts in the short run and overtime did not seem to provide support for previous research showing the effectiveness of salient recasts on learners’ learning. Loewen and Philp (2006), Egi (2007a), Sheen (2006), Kim and Han (2007), Nassaji (2007, 2009), and Han and Kim (2008) all regard saliency in recasts represented in different ways as paramount to favouring learners’ attention to the target form and learning. Different from Nassaji’s (2007, 2009) comparative study of the same type of feedback with different degrees of saliency, the present study did not find the significant differences between recasts with different saliency at any time. The explanations for these findings are fourfold: the present study’s stringent control measures, the combined use of another type of feedback in corrective recasts, learners’ intrinsic characteristics, and the different effects of normal and corrective recasts offsetting their respective advantage.
The present study considered the different variables possibly interfering with an unbiased evaluation of the effectiveness of recasts in its methodology. The situation of having the frequency variable fabricating the superior effect of prompts over recasts in those prompt-recast studies for instance would thereby be avoided. To prevent one feedback type from appearing more than the other type in any case, the present study tried to supply normal and corrective recasts consistently to every error committed by learners. As mentioned before, the Mann-Whitney U Test in chapter 5 illustrated the non-significant difference of recast attempts between the two recast types. The number of actions correspondingly obligating the similar number of verbs used by learners had been determined by the fixed story structure of the cartoon-strip narrative task. The number of learners’ use of verbs and errors could somehow be controlled. Moreover, both recast types were targeting one form, irregular past, throughout the feedback sessions. Their equal intensiveness could also be expected.

The saliency difference between normal and corrective recasts may have also been neutralized by the study’s attempt to make both recast types involve one error change only, to produce a sharp contrast between learners’ errors and the researcher’s recasts. Moreover, the lengths of both recasts were short by reducing learners’ utterances from entire ones to partial ones. The use of meaning-based narrative tasks and tests, with pre-task planning to ease learners’ online burden, to easily and appropriately elicit learners’ use of past tense may have further drawn the difference between learners from both recast groups closer.

As emphasized by Doughty (2001), consistent supply of feedback targeting a single predetermined form may prolong learners’ short attention window to form during online communication. However, same as Egi’s (2007a) study finding, the consistent and intensive supply of normal recasts may have raised their saliency level
to a level comparable to that of corrective recasts. In Erlam and Loewen’s (2010) recent study, explicit and implicit recasts were found insignificantly different in effecting learning, because of their equal intensiveness in repeatedly focusing on the same target form and implementation in a laboratory setting and foreign language context also involving form orientation. Furthermore, the communicative tasks designed to elicit the target form was found stronger than the feedbacks in facilitating learners’ use of form. Their findings may support the present study’s argument that the stringent controls of recast implementation and form-eliciting task may have neutralized the difference between normal and corrective recasts.

These control measures originally attempted to minimize the extreme difference between normal and corrective recasts. Their only difference can then be attributed to the saliency features of the preceding question prompts with stress on learners’ errors and the following stress emphasis on the target form in corrective recasts. However, these saliency features failed to contribute to the difference of corrective recasts from normal recasts. The reason could have been because these saliency features were in fact futile in making corrective recasts appear salient as intended. According to Sheen’s (2006) specification of features contributing to the explicitness of recasts, the present study’s implementation of corrective recasts violated two of the features stated. Firstly, the use of declarative mode to deliver recasts in a statement rather than in any discourse-like way such as interrogative confirmation check is advised; but corrective recasts were initially delivered in an interrogative tone to elicit learners’ self-correction. Secondly, the use of recasts in the form of full or partial repetition without any combined use with other feedback is recommended; but the question prompts preceding the subsequent recasts in corrective recasts appeared as one of the prompt feedback types—repetitions with
rising intonation and stress to highlight the errors (Lyster, 2004). These violations may have attenuated the intended saliency of corrective recasts.

No matter whether corrective recasts were more or less salient than normal recasts, their effects may have emerged the same to the present study’s learners, who may have either been too adapted to the formal orientation of both recasts or fossilized by their tenseless L1. According to Ellis et al. (2001) and Ellis (2009a), Asian students incline more towards form or knowledge focus due to their didactic education culture. Learners in the present study may have been equally receptive to both recasts. On the other hand, in line with Selinker (1972) and Swan (2005), the language development of the L2 exposure-poor learners in the present study may have been stabilized with staying at their tenseless L1 level, which remained equally difficult to be tackled by any kinds of consistent and salient feedbacks.

Even when the non-intrusive nature of normal recasts was suitable for learners’ receptive cognitive style, learners’ continuous mere reception and verbatim response to normal recasts may have lowered or discouraged their cognitive engagement to learn. This negative effect may have offset the non-intrusive advantage of normal recasts. Similarly, although corrective recasts bore additional saliency features to normal recasts in attracting learners’ attention, their intrusiveness to learners’ receptive cognitive style may have offset the explicitness advantage of corrective recasts. With these different effects offsetting the respective advantage of normal and corrective recasts at the same time, the difference between the two feedback types may have become less clear-cut.

The inconsistent quantitative results across feedback groups and time have been shown and explained by different factors in past research and the present study. Some factors are used universally to explain different findings. Other factors are used
respectively to explain different findings. Some factors however seem to contradict others. The stringent controls over the tasks/tests, predetermined target form, and procedures in the present study may explain why the feedback in normal recasts, though delivered directly with unclear negative evidence, captured and benefited learners’ immediate attention and learning. Meanwhile, the factor of stringent control suggests that the consistency, intensiveness and simple recast controls over normal recasts may have made normal recasts possess comparable saliency to that of corrective recasts, and contributed to the insignificant difference between them at any time.

Nevertheless, the factor of stringent control may create contradiction to the explanation for the insignificant difference between normal and corrective recasts, which suggests that the respective disadvantages of the two recasts offset their own advantages. The disadvantage of normal recasts in discouraging learners’ cognitive engagement could have been minimized by the study’s stringent control over the consistency and intensiveness of normal recasts. Normal recasts could have then borne the double advantages of being non-intrusive as well as stimulating learners’ cognitive engagement. For corrective recasts, the study’s equal stringent control over the consistency and intensiveness of corrective recasts could have made the originally explicit corrective recasts further salient to override its intrusiveness to learners’ receptive cognitive style. Corrective recasts could have then borne an enhanced degree of explicitness to stimulate learners’ cognitive engagement. Still, both normal and corrective recasts may have appeared equally advantageous to learners with the presence of stringent controls. Moreover, both normal and corrective recasts could have been equally non-intrusive when corrective recasts were in fact not as explicit as
what Sheen (2006) considers in terms of bearing interrogative discourse and a combined use with another feedback type.

The use of the explicitness in prompt-like corrective recasts to explain their immediate effectiveness may contradict with the ineffectiveness of the interrogative prompts in explicitly drawing learners' attention to form. The interrogative property has been argued as weakening the explicitness of recasts and falsely leading to learners' meaning focus. These may have resulted in the insignificant difference of corrective recasts from normal recasts and the non-lasting effectiveness of corrective recasts respectively. The contradiction can however be resolved by the possibility that learners' online processing of form had been expedited by the active negative evidence of the question prompts in corrective recasts. However, this fast activation of online processing may have only existed online or lasted not long after the online performance. As time passed, the online activation may have subsided, but the universal meaning focus of interrogative prompts may have lingered. Moreover, as the online activation had subsided overtime, the inherent disharmony between corrective recasts and the exemplar-based target form may have exerted more influence than before.

All in all, the stringent control of the present study, the suitability of the feedback for learners and the target form, and the implementation of feedback may have impacted the quantitative findings of the present study. Apart from the above possible justifications for the inconsistent quantitative findings and contradictions, the qualitative findings from learners' stimulated recalls may suggest some other underlying reasons for the results.
7.7 Stimulated Recall Theme—Task Demand

As analysed in chapter 6, some students in the stimulated recall subset revealed issues about the task demand cast on them at the time of the feedback sessions. Issues reported include forming the narrative story from bridging the meaning of different pictures, dealing with their unfamiliar experience with the narrative task, minding whether their narratives covered all the details required to convey the coherent and complete meaning, focusing on presenting narratives in their meaningful entirety, organizing their narratives for the purpose of meaningful expression and comprehension, and finishing their incomplete meaning previously interrupted by feedback. These issues were analysed as possibly leading to students’ unawareness of the feedback or the consistent use of past tense in their narratives. Even when they were aware of the feedback and the use of past tense, their prioritized attention was recalled having been placed on the above task demands.

These task demand issues raised by students provided support for previous research arguing that procedural demands of tasks, such as dual tasks and task unfamiliarity, may diminish learners’ capacity in attending to meaning and form simultaneously (Robinson, 2005, 2007; Skehan, 1998). On occasions where learners were able to attend to formal feedback under both the linguistic and procedural task demands, Robinson’s (2005, 2007) hypothesis that task complexity may push learners’ use of the target form as well as seeking help from corrective feedback may have been exemplified. Furthermore, VanPatten’s (1996, 2004) position may have been substantiated when learners recalled prioritizing meaning expression of the next pictures, content detail coverage of their narratives, meaning of the ending, and meaning organization. VanPatten upholds that learners tend to focus on meaning
more than form when attentional resources need to compete to process both meaning and form simultaneously.

However, some studies argue that as long as certain task designs and controls favour learners’ use of the target form and attention to the feedback, learners should have free capacity to defeat task demands obstructing their use and processing of the form. For instance, Ellis (2005a) specifies that the use of oral narration can often elicit learners’ obligatory use of the target form to construct story meaning. McDonough (2007) chose past tense as the study’s target form because of the ease of eliciting past tense in communicative tasks. Leeman (2007) also considers that learners’ form-meaning mapping can be facilitated by tasks with clear meaning. Not having to ponder on the task meaning, learners’ attention can be freed. Learners can thereby channel their attention to use form specifically for its corresponding meaning.

The present study had already structured its narrative task along a fixed storyline illustrated by pictures, showing some easy-to-identify actions. Moreover, Chinese summaries on the meaning of the cartoon-strips were given prior to the different tasks, to further familiarize learners with the meaning of the story. The narrative story had also been framed in the historical past, with the opening time prompt “long long time ago…” provided, to further elicit learners’ use of past tense. These control measures did not seem to have lessened learners’ online burden caused by the task demands, as unveiled by their stimulated recalls.

In other words, the online task demand burden on learners may have overridden the present study’s pre-task control measures intended to ease learners’ use of the form and attention to feedback. This may echo Ortega’s (1999, 2005) findings that her learners reported using the pre-task planning time to focus on the form to be used, but encountered difficulty with transferring their pre-task focus to
their online performance. Ortega attributed this to the overwhelming effect of the online task demand. The learners in her study revealed that they put priority on their online meaning expression, fearing that their listeners would be confounded with their incomprehensible meaning. Learners in the present study may have also worried about their incomprehensible meaning online. They may have hence placed more attention to bridging the sequential story parts, conquering their unfamiliarity with the narrative task, ensuring their coverage of the necessary content details, closing their narratives sensibly, and assembling the different idea units well to form comprehensible meaning. These online task demand issues may have exerted interference, regardless of the facilitation of procedural task controls and task complexity controls that Robinson (2005, 2007) and Skehan (1998) propose.

7.8 Stimulated Recall Theme—Speaking Modality

Another online challenge reported by students in their stimulated recalls was the speaking mode of the narrative task. It was found hampering learners’ meaning expression or progression, form usage and attention to feedback. Not having enough time to rehearse their thoughts and use of language, being occupied with keeping the continuous speaking flow, not enjoying careful offline planning as writing, having to face the L2 pronunciation challenge before conveying meaning and their use of grammar, and the passing duration of learners as well as interlocutors’ spoken messages were recalled as impeding learners’ attention to meaning, form and feedback during their online performance. These speaking constraints not only hindered learners’ narrative construction and form awareness, but also their determination to make changes and corrections when they were aware of form.
Learners recalled prioritizing speaking-specific conditions such as the immediate audience pressure and the unavoidable challenge of L2 pronunciation.

These learners' difficulties incurred by L2 speaking provided support for previous research identifying the limitations of speaking in promoting language learning compared to writing. Wolff (2000) contends that language learners learn more effectively from writing than speaking; because it allows a lasting medium and offline time for learners' clear, deep and controlled reception and processing of input and output. Empirical studies such as Wong (2001), Leow et al. (2008), and Niu (2009) confirmed that learners attended to form more in the written mode and had no problem attending simultaneously to both meaning and form written input. Especially in the case of L2 learners, more controlled processing is needed for them to attend to and perform L2 speaking because they lack the relevant procedural knowledge (De Bot, 1992; Ellis & Yuan, 2005).

Although writing and speaking bear their different inherent effects on learners' language learning, researchers such as Ellis and Yuan (2005), Ortega (2005), Sangarun (2005), and Yuan and Ellis (2003) posit that pre-task and online planning may compensate the lack of preparation room for learners' controlled processing in rapid online speaking. However, as Ellis (2006a) states, oral output entails automatic processing and thereby calls for learners' implicit knowledge; whereas written output involves controlled processing and elicits learners' explicit knowledge. It may be difficult for learners to employ their controlled processing allowed in the pre-task planning stage to facilitate their automatic processing in online speaking. The automatic processing demanded in online speaking is considered testing learners' feel for the use of form, rather than monitored use of form. Online planning may even
shape the speaking task entirely into a controlled task, calling for learners’ monitoring skills rather than automatized skills or communicative use of the form.

Eliciting learners’ feel for the form was exactly the present study’s purpose of using speaking tasks. Learners’ wrong feel for the form, due to their unproceduralized use, was expected to create opportunities for recasts to target; and the effectiveness of recasts can then be evaluated. The speaking mode and its inherent challenges may seem perfect in serving the study’s purpose. From the relevant stimulated recall data, the speaking mode was shown not only causing impact to learners’ use of form, but also their conception formulation. According to Leeman (2007), learners’ proceduralized use of form is often premised on their ease with meaning construction. In the face of difficulty with meaning construction, learners’ feel for the form may become a mission impossible.

As stated in Levelt’s (1989) speaking model, formulation of meaning precedes grammar usage, articulation of speech, and then monitoring of articulated speech. The obstruction to learners’ formulation of meaning by the online speaking constraints may accordingly impose problem to learners’ grammar usage, speech pronunciation, and awareness of their use of language. This chain effect caused did not seem to have been countered by the pre-task planning offered to learners in the present study. Learners’ performance may also be affected when the processing of any of these four stages gets obstructed. Levelt’s model applies to L1 speakers and hypothesizes that the four stages are processed automatically for L1 speakers. De Bot (1992) predicts that L2 speakers on the other hand process these four stages in a controlled manner. Obstruction caused to any particular stages or all stages by the online speaking constraints may have illustrated L2 learners’ controlled processing of the stages in Levelt’s model. The opposite automatized demand in the online speaking task seemed
to have overcome the controlled processing compensated in pre-task planning, as well as the controlled processing attempted by learners during the task.

7.9 Stimulated Recall Theme—Meaning and Form Competition

The narrative task in the present study was designed to elicit learners’ use of past tense to convey the historical and completed actions one after another along the sequential storyline. Learners’ attention to both meaning and form is thus crucial to their task accomplishment. Learners’ stimulated recalls suggested their different allocation of attention to meaning and form. For instance, some could afford to attend mainly to either meaning or form while managing to allot the rest of their attention to the other part. Some could only attend to either one at a time. Some other learners explicitly recounted their attention to either meaning or form at the expense of each other. Some reported their simultaneous focus on both meaning and form, but on the condition of transparent meaning. Some were able to attend to form as well as meaning, but the forms attended were non-target forms bearing more meaningful value than past tense. Some even perceived their tense mistake as pertaining to choice of meaningful vocabulary. Concerning meaningful value, some learners prioritized the meaningful value of their narratives to the immediate audience. They may have therefore chosen to use simpler meaning or accessible form to facilitate their strategic attention to meaning and form, in order to produce comprehensible narratives.

These findings about learners’ different attention allocation to meaning and form provided support for VanPatten’s (1996, 2004) hypotheses and studies. VanPatten has been one of the principal researchers investigating meaning-and-form competition for learners’ limited attentional resources. His Primacy of Meaning Principle argues that learners often prioritize meaning over form input comprehension.
when their attentional resources compete. *Meaning* is logically formed by content words which construct messages for learners to comprehend in order to attain communicative goals (VanPatten, 1996, p. 18). The sub-principle that “[learners] process content words in the input before anything else” (VanPatten, 1996, p.18) may be derived from VanPatten’s common-sense position that learners need to attend to and use content words to convey meaning. In view of this logic, the present study learner’s overriding focus on vocabulary search at that time may also exemplify VanPatten’s hypothesis of meaning and form competition. VanPatten (1996) defines form as “surface features of language” (p.18) that encode a message.

The input processing principles behind VanPatten’s processing instruction pedagogy propose that lexical or meaning-based attention, instead of syntactic or form-based attention, can be retained by learners under certain kind of task planning constraint. VanPatten’s input processing principles may help justify why some learners of the present study had, as shown in their stimulated recalls, de-emphasized the use of the form (past tense) expected in the spoken narrative task, when their cognitive resources had been constrained by the content-based pre-task planning. VanPatten’s (1996) central principle that learners “process input for meaning before they process it for form” (p.17) may conclude learners’ practice of primarily focusing on meaning over form.

VanPatten makes use of both first and second language researchers’ common view that language acquisition requires attention to the target input as the point of departure for his input processing principles. For example, Slobin (1985) contends that attended input determines first language acquisition; and Schmidt (1990) argues in a similar vein that second language acquisition builds on attention to formal input. These researchers hypothesize that attention is the key to language acquisition;
however, other researchers (e.g. Skehan, 1998) suggest that the human information processing system involves limited attentional capacity, in handling for example multiple input sources or heavy computation load at a time. VanPatten notes the likely side effect of the detection process proposed by Tomlin and Villa (1994)—learners’ detecting one input source will interfere with their detecting other input sources. VanPatten (1996) puts forward two overarching input processing principles to outline the input interference phenomenon. The first one is the aforementioned central principle; and the second is, “[for] learners to process form that is not meaningful, they must be able to process informational or communicative content at no or little cost to attentional resources” (p.27).

Even when attention is exclusively allocated to form, the Lexical Preference Principle (VanPatten, 1996; Benati, 2005) predicts that learners prefer to process lexical items to grammatical morphology sharing the same meaning. Benati (2005) found that his Chinese learners used temporal adverbs, also because of their L1 transfer, to unequivocally replace tense and aspect in English for the past time meaning. This preference of learners may have turned past tense redundant, although past tense has both its meaningful and grammatical functions favouring form-meaning mapping (Leow et al, 2008).

When input comprehension concerns learners’ perception of the focus of the feedback, Mackey et al.’s (2000) study showed that lexical and phonological feedbacks were perceived correctly more than morphosyntactic feedbacks. They attributed this finding to the contribution of lexical and phonological focuses to learners’ understanding of the ongoing communication; whereas morphosyntactic focus was considered less functional in this regard because of its low communicative
value in motivating learners’ noticing in a meaning-oriented context (VanPatten, 1990, 1996).

A representative empirical study conducted by VanPatten himself in 1990 may substantiate his own principle that meaning or lexis wins learners’ attention over form. It may also validate the stimulated recall findings about learners’ primary focus on meaning and lexis, among focuses of the written Chinese summary input, pictorial input, formal feedback input and speech output processing of the story. There were three experimental groups and one control group in VanPatten’s (1990) study, to investigate whether English learners’ meaning comprehension of a passage in Spanish would be affected by their simultaneous focus on the use of a content word, an article, and a verbal inflection. The results revealed that learners’ attention to the content word designated did not cause any interference to their meaning comprehension of the passage. They were able to recall as many meaning idea units of the passage as the control group, which only focused on meaning comprehension of the passage. The other two experimental groups, which were asked to focus on grammatical form while attending to meaning of the passage, performed significantly worse than the content word and control groups in recalling meaning idea units of the passage.

Since VanPatten focuses on learners’ unidirectional comprehension of aural input, some may argue that his theories are irrelevant to the present study, which incorporated input and output processing. However, the main focus of the feedback session or stimulated recall session was learners’ perception of the aural feedbacks, which involved learners’ comprehension of aural input. The present study overall comprised written input from the story summary in Chinese provided in pre-task planning, pictorial input from the cartoon-strip pictures as visual aid during task, formal feedback aural input via recasts during task, and learners’ narrative speech
output. Apart from this, VanPatten’s principle is also relevant to the present study’s qualitative findings based on its underlying hypothesis that learners’ attentional resources are limited; and the resulting taxing cognitive processing leads to the competition between meaning and form processing.

In addition to explaining learners’ biased attention to meaning and form, VanPatten also suggests that saliency added to input may resolve the problem and facilitate learners’ simultaneous processing of both meaning and form. Accordingly, the consistency and intensiveness in the two recast types in the present study as well as the additional saliency in corrective recasts should have evened all learners’ attention to meaning and form. Moreover, as Ellis (2009a) notes, meaning-based tasks specially designed to elicit the use of certain form together with consistent and rich input feedback as scaffolding may facilitate learners’ simultaneous communicative and formal focuses. As recalled by learners in the present study, only some of them were able to simultaneously attend to meaning and form, while others were unable to achieve that.

This could have been due to the present study EFL learners’ inherent formal deficiency or insensitivity, and lack of speaking practice. Skehan and Foster (2001) correlate L2 learners’ formal deficiency or insensitivity to their limited attentional capacity to attend to both meaning and form. Learners often as a result resort to communicative strategies to attend to and produce meaningful units only for communicative purposes. With their continuous negligence or insensitivity to form, learners’ interlanguage development is easily stabilized. Swan (2005) on the other hand relates learners’ formal deficiency or insensitivity to their limited exposure to the target language in EFL classes. Learners’ formal foundation consequently becomes inadequate for performing communicative tasks. Based on the findings of
these researchers, the inherent formal deficiency or insensitivity of the EFL learners in the present study may have made them immune to the facilitative intervention of salient feedbacks and meaning-form-mapping tasks.

The EFL learners in the present study were rather foreign to online L2 speaking, because teachers often scaffold L2 meaning expression for them to avoid overrunning their short lessons (Philp & Tognini, 2009). In the light of the discussion of section 7.8, the present study’s use of the speaking mode, which entails planning, time, and memory limitations as words are spoken immediately as they are decided at that moment (Bygate, 1987), may have further imposed online cognitive pressure on learners. These could have been the reasons why some learners encountered the cognitive competition between meaning and form processing.

Nevertheless, when meaning was transparent to learners or made simple by learners, as articulated by some learners in their stimulated recalls, the online constraint of speaking and their inherent formal deficiency did not seem to have interfered with learners’ simultaneous attention to formal feedback. This suggested that any one factor which helped relieve learners’ cognitive burden may have facilitated their simultaneous attention to meaning and formal feedback, regardless of other concurrent constraints. Opposite to the phenomenon of meaning overriding form hypothesized by VanPatten, some learners reported their primary focus on form over meaning. This may have been due to their formal deficiency of automatizing formal usage to map with meaning, and thus dwelling on their declarative knowledge to inform formal usage to fulfil meaning in the tasks. Learners’ unproceduralized use of form is what the next section will turn to.
As revealed in their stimulated recalls, the language development of the present study’s learners was staying at their interlanguage and unproceduralized L2 stages. They reflected that they were not aware of the need to use past tense or the stable use of past tense for their narratives until the intervention of recasts. Some of them reported the same unawareness despite their actual use of past tense in the task. These suggested that learners were still developing the tensed L2 from their tenseless L1. Other feedback and recall excerpts illustrated that students boldly adopted their L1 grammar of legitimately dropping verbs; and some of them were aware of that, while others were not. Some learners’ spontaneous tense, tenseless and verbless use of grammar had been recollected as also due to their distracted attention to meaning and online speaking constraints.

Some other students were however aware of the use of past tense and were thinking of being careful about it. Their problem exhibited was with actually applying the use of past tense to their narratives. Some were able to apply the use of past tense but not consistently throughout their narratives. These were all occasions indicating learners’ unproceduralized L2. Learners were relying more on declarative knowledge to inform their use of L2 in other words. This could have led to learners’ mindfulness over grammar and their own mistakes. Some students extraordinarily attributed their selective use of past tense to their comfort level with certain verbs. This also suggested learners’ unproceduralized use of L2 across verbs.

The findings about learners’ interlanguage and unproceduralized use of past tense provided support for past studies arguing that learners’ proficiency involves different spectrums. Ellis (2006a) has broadened the understanding of learners’ language knowledge by viewing the difficulty of each grammar form according to
their explicit and implicit knowledge representations to learners. For example, Ellis reckons past tense as both easy and difficult to learners. He considers that regular past is easy as explicit knowledge because of its transparent rule, but difficult as implicit knowledge in actual use. Irregular past is on the other hand regarded easy as implicit knowledge because of its item-based property, but difficult as explicit knowledge without a general rule. Rothman (2007) particularly explains the different spectrums of language knowledge with the case of Chinese learners. Chinese learners were depicted as easily grasping L2 knowledge, but unskilful at performing L2 knowledge in actual production. These researchers precisely elucidate the present study learners’ relative readiness of possessing declarative past tense knowledge but unproceduralized past tense knowledge (Anderson, 1983).

According to Anderson’s (1983) early proposal that declarative knowledge can be automatized via practices and turned ultimately to proceduralized knowledge, the present study learners should have been able to display their proceduralized use of past tense with the consistent, intensive and salient practice of receiving and repeating recasts. With both normal and corrective recasts being near-equally explicit, learners should have received straightforward enough reminders of their misuse of tense (Scheffler, 2008), and displayed automatic use of past tense after initial reminders. Especially when learners were developmentally ready for past tense and it was not entirely novel to them, they should have benefited from initial recasts and displayed consistent use of past tense.

However, one point needs to be raised is that the feedback session was never a prolonged one extending across time. Therefore, learners’ display of proceduralized use of past tense may not have been available until a later time, perhaps in the delayed post-test. The correlation between the stimulated recall and delayed post-test findings
Another point is that proceduralized use of past tense relates to implicit knowledge exhibited in learners' actual use or feel for the form in a task; and implicit knowledge of past tense often takes time to develop (Bardovi-Harlig, 2000). For these reasons, the different facilitative measures of past tense-eliciting task, developmentally ready target form, and explicit recasts offered within the rather short treatment time may not have been effective in immediately driving learners' show of proceduralized use of past tense.

Even though it is possible for learners to exhibit proceduralized use of past tense within the short feedback session due to the above facilitative measures, learners' reception of the help from recasts may not be guaranteed. Yoshida's (2008, 2010) learners showed surface signs of receiving and understanding recasts, but they confided that they did not notice the contrast drawn between their errors and the target form. They revealed their contemplation over their own errors and sentence construction at that time. They also verbalized their preference for feedbacks eliciting correction because they could thereby gain more time to think. Yoshida's learners were even strategically being agreeable to teachers' recasts so as to avoid online embarrassment. Learners' intent other than learning from recasts is difficult to control. Thus, any online or immediate learning signs may be false or even impossible.

7.11 Stimulated Recall Theme—Different Effects of Corrective Feedbacks and Uptakes

Learners' stimulated recalls advised different effects of recasts and uptakes which may have influenced learners' learning from recasts. Some learners noticed the corrective purpose of recasts and the use of past tense from previous recasts. The effect of previous recasts had been extensive across turns, although they were
targeting their respective erroneous episodes. The regularity of recasts had also been shown helpful to facilitating learners’ awareness of the corrections and functions of recasts by their recurrent presence and reminders. Some learners contrastively perceived recasts negatively and criticized them as interrupting their other concurrent thoughts. Other learners clearly recalled their unawareness of the formal focus of recasts and treated recasts as merely serving communicative confirmations.

Another recast-related issue unveiled by learners’ stimulated recalls was their perception of uptakes. Some learners treated their immediate uptakes of recasts as practice opportunities to train their current as well as subsequent use of past tense. Moreover, some learners used uptakes not only to confirm their meaning expression, but also their awareness of the correct use of tense to the researcher. Besides these effects of uptakes facilitating and showing learners’ formal awareness, learners also perceived uptakes in ways unrelated to formal awareness, such as employing uptakes to help them transit to the next narrative meaning. Some other students however employed no uptake but subsequent use of past tense in other instances to show their formal awareness. They attributed their zero use of uptakes to dealing with the ongoing speaking flow and the redundancy to repeat the same mistake when it had already been noted.

These findings about the effects of recasts and uptakes and learners’ perception towards them provided support to studies proposing and demonstrating the relevant positive and negative effects. The extensive effect of feedback can be justified by Ellis (2006b) and N. Ellis’s (2005, 2006) proposal of the interface between explicit and implicit knowledge. Taking a more moderate approach, Ellis (2006b) considers that explicit knowledge will subsequently turn to implicit knowledge on the conditions that learners are developmentally ready and explicitly drawn to the target
form by intensive feedback. N. Ellis (2005, 2006) strongly views that frequent and salient exposure to recasts and practice alone will facilitate the interface. The explicit knowledge delivered by the frequent and salient recasts, and its extensive facilitative effect on learners' subsequent automatic usage of past tense (implicit knowledge) in the present study, may reflect the two researchers' hypotheses.

The facilitative effect of recasts shown from its consistency may strengthen Han (2002) and N. Ellis's (2005, 2006) position that the consistent and frequent use of recasts shapes the effectiveness of recasts on learners' learning. The downside of the consistent supply of recasts, though it fostered learners' awareness, was making learners feel irritated as raised by Truscott (1999, 2007). The communicative function of recasts interpreted by learners echoed what the L2 classroom studies discussed in section 2.3 of chapter 2 found of the ambiguity of recasts incurred by their dual functions as communicative and corrective. Regarding the practising value of uptakes as recognized by learners, DeKeyser's (2001, 2007) stance that learners' repetition of recasts helps them proceduralize the knowledge has been supported. The credit of uptakes has been further demonstrated by learners' recalls of using uptake to confirm their formal as well as meaning awareness. This finding coincides with Lyster and Ranta (1997) and Yoshida's (2010) acknowledgement of uptakes as functional in suggesting learners' instantaneous noticing and learning, but not necessarily representing learners' instances of learning. Learners' use of uptakes for other purposes and unnecessary use of uptakes to indicate their noticing and learning had also attested the indicative rather than affirmative functions of uptakes.

The findings about learners' different perceptions of the effectiveness of recasts and use of uptakes may raise questions about studies suggesting the categorical effectiveness of recasts and uptakes within individual studies. With the
control of variables, such as feedback consistency and learners’ developmental readiness, both past studies and the present study’s quantitative findings illustrated certain categorical patterns of the effectiveness of recasts. However, with the insightful qualitative investigation of learners’ stimulated recalls, different patterns of the effectiveness of recasts and the use of uptakes have been revealed. This discovery corroborated Leeman (2007), Egi (2007b), and Yoshida’s (2010) belief of the usefulness of using learners’ perception to gain insights into phenomena unobtainable through quantitative research.

The present study therefore incorporated both quantitative and qualitative methods to examine the effectiveness of recasts from general as well as exhaustive angles. It adopted Polio et al.’s (2006) approach to use stimulated recalls to redress the limitation of quantitative investigation in probing the variable of different individuals. The variable of individual participants had been controlled by the present study’s measure of distributing participants equally from different classes across different feedback groups. This was only to avoid skewed quantitative results, but never meant to eliminate the variable of individual differences. The emergence of different learners’ perception of recasts and uptakes may enlighten the different reasons behind the effectiveness and ineffectiveness of recasts and uptakes.

7.12 Correspondences and Differences between Quantitative and Qualitative Results

Comparing the two sets of data, the stimulated recall findings that recasts projected extensive and consistent effects to learners’ formal awareness and uptakes carried proceduralizing functions appeared to correspond to the positive findings of hypotheses 1 and 3. The specific positive effects of both recasts and learners’ uptakes all happened within the feedback session. The immediate effectiveness of both normal
and corrective recasts that the quantitative results found may have been derived from the facilitative influence of recasts and uptakes produced recently in the feedback session.

The online issues of task demand and speaking modality, the exposure issues of learners' formal insensitiveness and speaking practice deficiency, and the learner issue of developing language mastery may explain the insignificant findings of the lasting effect of corrective recasts and the immediate and lasting difference between normal and corrective recasts. The negative effects of recasts, and learners' irrelevant and absent use of uptakes found in the stimulated recall findings may also be the reasons behind the aforementioned insignificant findings.

A mismatch between the two sets of data however occurred when the sustainable effect of normal recasts stood, despite all the issues or constraints revealed by learners in their stimulated recalls. The intrusive effect of recasts found in learners' stimulated recalls may have been less influential in the case of normal recasts than corrective recasts. As discussed in section 7.3, the directness of normal recasts seemed to match with Chinese learners' passive cognitive style. The effectiveness of normal recasts may have thereby sustained. Some learners who did not initiate any verbal uptakes, no matter whether they noticed recasts or not, may have also been affected by their passive cognitive style.

7.13 Conclusion

Prominent issues that have arisen in the discussion of the quantitative findings are the weakening of variable differences by a well-controlled study, the harmony between the feedback and learners' cognitive style, and the questionable explicit implementation of corrective recasts. For the qualitative findings, the main issues to
have arisen are online constraints, learners' exposure limitation and their developing
mastery. These issues did not seem to have been widely correlated to the effectiveness
of recasts in the field. Different perceptions of the advantages and disadvantages of
recasts have been revealed by different individuals, despite the generic effects of
recasts found in the quantitative analysis. This may illuminate the complementary
value of using a qualitative approach. The use of stimulated recalls to bring learners' awareness or perception to the surface however remains questionable. Learners' recalls may have been their new thoughts created during the interviews. Moreover, the different schedules of learners' participation and the consistent focus on past tense may have enhanced learners' recalls of their awareness of past tense, through communicating with other participants and the reinforcing effect of recurrent past tense focus respectively. The above arguments explaining the study's quantitative and qualitative findings are suggestive rather than definitive.
Chapter 8 Conclusion

8.1 Thesis Summary

This thesis has documented the present study's investigation into the respective and comparative effects of enhanced and unenhanced recasts on Hong Kong secondary school learners' use of irregular past in their spoken narratives. Apart from measuring the feedback group difference among learners, the test time difference among them was also probed to gauge their learning development. Since the controversy over the effectiveness of recasts lies in learners' subsequent learning evidence as well as their noticing of recasts, the qualitative method of stimulated recalls was used to examine learners' perception towards recasts and their use of past tense during the feedback sessions. Different control measures were implemented in the study's methodology to keep any interference to the quantitative and qualitative results to the minimum. The control measures used were based on a thorough review of the past literature on recasts and second/foreign language acquisition. The two data sets provided room for the present study to view the effectiveness of recasts in both global and in-depth levels. Issues correlating the findings of the two data sets were discussed according to past studies and the present study.

The quantitative results of the present study found that both normal recasts (unenhanced) and corrective recasts (enhanced) led to learners' immediate gain in using past tense in their spoken narratives. Their lasting gain was however sustained by normal recasts but not by corrective recasts. The hypothesized superior effect of corrective recasts over normal recasts was not confirmed in the short run as well as overtime. The qualitative results from learners' stimulated recalls found that different underlying issues revealed through learners' perception of recasts and their use of past tense during the feedback sessions may have governed learners' later performance.
Issues such as task demand, speaking modality, meaning and form attentional competition, learners’ language development status quo, and different effects of recasts and uptakes to learners were discovered possibly affecting learners’ reception of recasts and subsequent learning.

The significance of these results was that enhanced recasts were not necessarily facilitative to Hong Kong learners’ use of past tense in their spoken narratives. The factor of saliency was also found not necessarily contributing to the difference between normal and corrective recasts. The sustainability of the facilitative and hence learning effect of recasts may not have depended on saliency. The unexpected findings of the lasting effect of normal and corrective recasts and their insignificant difference across time may have been due to the different online task and speaking constraints, learners’ inherent language deficiency and cognitive style, and learners’ different perception of recasts and uptake. Both the quantitative and qualitative results pointed out the divergence from what past studies commonly found. In other words, saliency does not statically or categorically lead to learners’ learning improvement and sustainability. Learners’ perception of the saliency in recasts and the concurrent interference incurred by the online cognitive constraints may have governed the effectiveness of salient recasts.

8.2 Discussion of Findings

The different possible explanations for the quantitative and qualitative findings and the correlation between the two sets of findings have been provided in the last chapter. With reference to the literature and the present study, the immediate gain of the normal and corrective recast groups may have been due to the stringent study control and explicitness respectively. The lasting gain of the normal recast
group was speculated to be derived from the non-intrusiveness to learners’ cognitive style and skill development. On the other hand, the intrusiveness of corrective recasts may have led to learners’ non-lasting gain from corrective recasts. The problematic interrogative feature of the question prompt in corrective recasts may have attenuated the explicitness, causing both the non-lasting effect of corrective recasts and the insignificant difference between the two recast groups. The controlled delivery of both recast types possibly neutralized their hypothesized difference.

Overall, the present study’s stringent control of the implementation of recasts, learners’ cognitive style, the problematic explicitness in corrective recasts, and the non-intrusive nature of normal recasts may have led to the inconsistent quantitative results. There were however some less firm explanations, such as the consistent supply of explicit recasts may have led to the effectiveness of corrective recasts. Meanwhile, learners’ effortless reliance on the consistent explicitness may have weakened their independence to use the target form at a later time when explicit recasts were no longer provided. Another less certain explanation was whether it was learners’ entrenched L1 knowledge or their habitual reception of formal input that made them unable to benefit from the two recast types differently.

8.3 Implications of the Study

8.3.1 Implications for Research

The implications of the study can be divided into three aspects: research, practice, and policy. The research aspect involves both theoretical and methodological concerns. Regarding the theoretical implication of the present study, the positive evidence in recasts may excel the negative evidence in recasts in effecting learners’ learning of the target form in the short as well as long run, from the long-lasting effect
found from normal recasts. This was particularly the case with the Hong Kong learners in the present study, who were used to receptive learning of receiving teachers’ instruction unidirectionally with little students’ involvement.

Moreover, recasts may not be absolutely implicit based on their communicative congruity with the ongoing interaction. Whether or not recasts are explicit or implicit depends on how they are implemented in a classroom or experiment. Recasts can be as explicit and correction-eliciting as prompts by adding preceding question prompts or repetitions. However, the prompt-like feature added to recasts may not guarantee learners’ benefit, because its intrusiveness to the ongoing flow and interrogative tone may interrupt and distract learners’ formal focus respectively. Furthermore, EFL learners’ formal orientation or lack of practice to respond to recasts may not obstruct learners’ learning from communicative recasts; because it all depends on how formal or communicative one wants recasts to be. The theoretical implication brought by the present study is that recasts are versatile. They can be in any degrees of explicitness; in other words, they can be made as effective as one would want them to be.

Nevertheless, from the present study’s methodological control of the implementation of normal and corrective recasts, measures intended to maximize the explicitness of recasts in a study may turn out to be obscuring results. The near-equal consistency and intensiveness of the two recast types, originally intended to isolate the variable of tonal and prompt-like saliency in corrective recasts for accurate comparison, emerged as moderating the difference between normal and corrective recasts. The saliency of normal recasts may have been pushed to a comparable level to that of corrective recasts. Another methodological implication is that task
procedural controls may not overcome the immediate and hence overwhelming interference caused by online task demands as well as learners' cognitive constraints.

8.3.2 Implications for Practice

Pedagogically, teachers may need to consider the suitability of certain types of feedback for their learners' cognitive style and the particular target form. Although giving consistent personal feedback intensively on a single form across time may enhance learners' reception and learning, intrusive feedback to learners' habitual learning and inappropriate feedback for certain form may largely ruin the effectiveness of the feedback. Therefore, teachers may want to relate the type of feedback to be used to their learners and the target form before making any predetermined error correction decisions.

8.3.3 Implications for Policy

To encourage teachers' use of one-on-one corrective feedback and suitable corrective feedback to facilitate learners' use of form in achieving communicative goals, educational authorities in Hong Kong may need to expedite the implementation of small-class teaching. With smaller class sizes, teachers may have more time to allow learners' experimenting their form usage in communicative activities and teachers' corresponding personal guidance. The exam-oriented school practice should also be reduced or changed to communicative assessment. This may thereby offload teachers' work and permit them time and concentration to study which corrective feedback would be suitable for their learners and the target form in a certain lesson or activity.
8.4 Limitations of the Study

Notwithstanding the present study’s effort on variable control, it involved some limitations which need to be considered in future research. The study was a quasi-experimental design. Such a design does not entail a randomized selection of participants. The non-randomized participants chosen within the same target school and English proficiency stream may have biased the results. This was because the medium proficiency stream in the target school could have been different from the medium stream in other schools.

Speaking of participants, the sample size of the present study may not be large enough to represent the F.3 secondary school population in Hong Kong. The total number of participants was 89, and results obtained from these 89 students from the same target school may only be the case of that particular school. Even the results represented the target school, they may not be generalized to the classroom practice of the target school because of the experimental setting of the study. The purpose of using an experimental setting was to control extraneous variables, but the downside was that the results may not extend to the classroom reality and advise teachers’ practice.

One widely discussed methodological limitation was the use of stimulated recall to push learners’ inner thoughts to a surface level through verbalization. Although learners’ recalls of what they were thinking of the recasts and their use of past tense during the feedback sessions were suggestive of their awareness, it remains doubtful if the recalls accurately reflected their cognitive states at that time. The delay in recalling may have corrupted learners’ memory; and learners may have created new thoughts just to fulfil the interview requirement.
The narrative task used in the study may have affected learners’ performance. The study chose to use a cartoon-strip narrative task because of the suggestions given in the literature that it can fix the meaning to free learners’ cognitive burden and elicit learners’ use of past tense. However, learners’ unfamiliarity with the task type and online speaking mode may have obstructed their awareness of formal feedback, use of past tense and construction of meaning.

Learners’ use of irregular past was measured through their accurate forms, partially correct forms, and target-like use. However, their overuse of certain forms, such as was and go, may have been due to their familiarity with or quick memory retrieval of these two tokens instead of their genuine use of past tense. The possibility of overuse was not taken into account in the analyses.

Based on the above limitations reported, future research may try to adopt an experimental design involving pure mixture of different participants from different schools. This may also help enlarge the sample size to make results more representative. Moreover, future studies may invite teachers’ participation to carry out controlled recast delivery in their classrooms, to reflect what teachers can do in reality as much as possible. For the use of verbal protocols to examine learners’ perception, future studies may incorporate a quantitative approach to analyse learners’ perception, like Mackey et al.’s (2000) study, to confirm the qualitative findings of learners’ perception. To familiarize learners with the narrative task type and speaking mode, future studies may invite teachers’ participation as well as request their adoption of the study’s task type and speaking mode in their regular lessons beforehand. A more valid picture of learners’ interlanguage change in using target construction would have been obtained if their overuse of certain forms had also been considered, to comprehensively and closely study indices of their interlanguage use. These
suggestions may help refine future research on similar topics; but the controls on learners, the target form, the task, and procedural variables, as implemented in the present study, are still recommended to secure research reliability and validity.

8.5 Overall Significance

The overall significance of the present study is that it contributed to the understanding of the changeable effectiveness of recasts and the interference to the effectiveness of recasts by online cognitive demands. Past studies have commonly attributed the ineffectiveness of recasts to their implicitness of integrating form focus with the ongoing communication, and hence concluded that recasts are ambiguous. As some past experimental studies found, the effectiveness of recasts can be facilitated by consistency, intensiveness, or saliency. The present study on the other hand extended this current knowledge about recasts by suggesting that consistency and intensiveness may have been sufficient to make recasts explicit and effective. This conclusion is developed from the insignificant difference between normal and corrective recasts across time. The near-equal amount and degree of consistency and intensiveness may have already pushed non-salient recasts (normal recasts), in the form of no tonal change and preceding prompts, to a similar level to their salient counterpart (corrective recasts). The tonal change and preceding prompts in corrective recasts may have been intrusive and hence unhelpful to learners. The direct reformulation in normal recasts may have contrastively suited Hong Kong learners’ receptive cognitive style.

Apart from enlightening the changeable effectiveness of non-salient recasts and their suitability for Hong Kong learners, the present study also suggested online cognitive burden as the possible cause of the ineffectiveness of recasts. Task
demands, speaking modality, attentional competition for form and meaning, learners’ unproceduralized use of language in the actual task, and learners’ different perception towards recasts and uptake during the task may have contributed to learners’ online cognitive burden. Few studies have drawn the correlation between these online issues and the effectiveness of recasts.

The results of the sufficiency and appropriateness of consistent, intensive, and non-intrusive normal recasts in effecting Hong Kong learners’ past tense learning may further shed light on the applicability of the cognitive-interactionist SLA approach to the Hong Kong context. Although Hong Kong learners may not be familiar with receiving and responding to personal recasts, their attention to and hence sustainable benefit from recasts can be facilitated by controlling the implementation of recasts. Moreover, the directness and non-intrusiveness of normal recasts may have been shown compensated Hong Kong learners’ unfamiliarity with proactive feedback delivery, and thereby facilitated their learning of using the form for communicative meaning. Even with the well-controlled implementation and cognitive suitability of normal recasts for Hong Kong learners, the on-the-spot cognitive constraints occurred in the task may have overwhelmed the tailor-made harness of delivering controlled and non-intrusive normal recasts for Hong Kong learners.

8.6 Reflections on the Research

Throughout the process of conducting the present research, I have learnt that implementing stringent control of variables in the methodology may not definitely give rise to fair results. The stringent control turned out to affect the comparison between normal and corrective recasts. Moreover, taking reference to the literature in advising the use of narratives to elicit learners’ use of past tense may not be sufficient.
Instead, taking learners’ familiarity with narratives into consideration emerged to be essential when designing tasks or test instruments. Task familiarity may lessen learners’ online cognitive burden. If I were to do the research all over again, I would either adopt task types which teachers usually do with students in their regular lessons, or request teachers to adopt the study’s task type in their regular lessons beforehand to familiarize students with it. More importantly, I would invite teachers’ participation to allow result generalization from the experimental study to the classroom reality.

However, the complexity of investigating the effectiveness of recasts in second or foreign language acquisition is that spontaneous delivery of recasts may conceal the potential effectiveness of recasts, when recasts can be in explicit forms like other explicit feedbacks shown to be more effective than recasts in past studies. Moreover, researching recasts in non-communicative pedagogical contexts may require a certain degree of controlled adaptation of recast implementation, to give a context-specific view of their effectiveness. A context-specific view of the effectiveness of recasts may be necessary since recasts might be taken up very differently in different language or cultures. With recasts originated from children’s first language acquisition in North America, Asian students, who are more reticent than Western students in their teacher-dominant education culture, may not be able to perceive the correctiveness of communicative recasts and show their immediate uptake of recasts. This cultural element needs to be explored in future research.

In terms of exploring the feasibility of using recasts as a pedagogical innovation in Hong Kong to facilitate learners’ communicative use of grammar, the use of communicative tasks, though unfamiliar to Hong Kong classroom learners, may be an undeniably advisable instrument to test the future use of recasts for training.
learners’ communicative use of past tense. Recasts could be a pedagogical innovation and success in Hong Kong based on the initial prediction of the present study that Hong Kong students have acquired the declarative knowledge of grammar form but not the proceduralized knowledge. However, students who do not have both kinds of knowledge, like students in Japan who learn English at a later age than students in Hong Kong, may be benefited more from explicit teaching of rules. The effectiveness of recasts in contexts like the Japanese education context is also needed to be explored in future research.
References


Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 197-261). Cambridge: Cambridge University Press.


Ellis, R. & Yuan, F. (2005). The effects of careful within-task planning on oral and


Ransdell (Eds.), *The Science of Writing*. Mahwah NJ: Lawrence Erlbaum.


Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation


Sheen, Y. (2006). Exploring the relationship between characteristics of recasts and


Stibbard, R. (2004). The spoken English of Hong Kong: A study of co-occurring...


122.


