Focus on Form in Task-based Language Teaching

- Exploring the Effects of Post-task Activities and Task Practice on Learners' Oral Performance

LI, Qian

A Thesis Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

in

English (Applied English Linguistics)

The Chinese University of Hong Kong

July 2010

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Thesis Committee

Professor Jette G. Hansen Edwards (Chair)

Professor Peter Skehan (Thesis Supervisor)

Professor Gerald Nelson (Committee Member)

Professor Pauline Foster (External Examiner)

Declaration

1. The Thesis submitted herein is for the fulfillment of the graduation requirement of the degree and programme stated.

2. I declare that the Thesis submitted is my original work except for source material explicitly acknowledged. I also acknowledge that I am aware of the University policy and regulations on honesty in academic work, and of the disciplinary guidelines and procedures applicable to breaches of such policy and regulations, as contained in the website http://www.cuhk.edu.hk/policy/academichonesty/.

3. I hereby agreed to grant the University a worldwide irrevocable, non exclusive right in respect of the copyright in the Thesis for the purpose of making copies, abstracts, reproducing or otherwise dealing with the Thesis by whatever means, including but not limited to, digitizing, storing, reproducing, and distributing the Thesis in any media and in any format, provided that any and all such acts are only for scholastic and academic purposes and with proper acknowledgment of authorship. This right is for the term of the existence of copyright in the Thesis.

Signature Li Qian Date July, 2010

Abstract

Li Qian

Given the small body of existing research concerning focus on form at post-task stage in task-based language teaching, the present study adopts post-task transcribing as a focus on form activity and explores the effects of transcribing under various conditions. In addition, two task types are adopted in multiple task sessions to explore task effects and task practice effects on learners' oral performance.

Ninety-six participants, divided into five experimental groups and one control group completed four tasks with a one-week interval between each task. Different experimental groups were assigned various post-task activities respectively. No post-task activity was adopted in the control group. Task performance was measured in terms of complexity, accuracy, fluency and lexical performance.

The findings are multifaceted. First of all, the adoption of post-task transcribing was found to be efficient for different formal aspects of task performance. In the second place, the pair-based transcribing brought about more syntactically complicated language, whereas the individual condition at the post-task stage led to an improvement in lexical sophistication. Thirdly, further revision after transcribing had complex effects on accuracy and complexity. Fourthly, interactive tasks proved to be more promising for a better overall task performance. Last but not least, multiple task practices were found to be beneficial for learners' lexical performance.

The findings were discussed in light of the concept of noticing and attention,

Levelt's speaking model, socio-cultural theory and other related SLA theories. Based on the theoretical discussion, pedagogical implications have been proposed.

This research argues that in task-based language teaching, more attention should be paid to the post-task stage regarding its effect on focus on form. Specifically, it suggests that a) a post-task transcribing activity can be adopted as a feasible focus on form activity in L2 classrooms; b) different conditions for the operation of post-task transcribing may bring about distinct effects on various aspects of task performance; c) different task types have different effects on learners' performance; d) multiple task sessions are necessary for L2 language improvement. Further, the present study calls for a process-product approach in further studies concerning the effects of post-task focus on form activities.

摘 要

鉴于在英语任务型教学研究中对任务后阶段研究的匮乏，本研究尝试在任务后阶段让受试对象进行聚焦于语言形式的口语书面抄录活动，探索在各种不同条件下进行的书面抄录活动对二语学习者任务表现的影响。另外，受试对象在四次实验小节中分别完成了两种类型的任务。因此，本研究同时探讨多次任务练习和不同任务类型对任务表现的影响。

96 名受试被分至五个实验组和一个控制组。六组受试均完成四个任务（两个讲述任务和两个决策任务），每次任务小节间隔一周时间。不同的实验组在不同的条件下进行任务后的书面抄录活动。控制组无需进行任何形式的任务后活动。受试的任务表现从语言的复杂度、准确度、流利度和词汇表现等方面衡量。

本研究的发现涵盖了以下各方面。首先，实验结果证明，在任务后阶段进行书面抄录活动有利于提高学习者进行任务时的语言表现。其次，各种不同的书面抄录条件对任务表现的影响也不一样。两人成对进行的任务后抄录促使学习者运用句法复杂度更高的语言；而与之相反的是，个人单独进行的任务后抄录则促使学习者运用更加复杂的词汇。另外，部分受试有机会在书面抄录后对自己的书面任务进行进一步的修改。这样的条件提高了他们的语言准确度，却降低了语言复杂度。与单独进行的讲述任务相比，学习者在成对交流的决策任务中的语言复杂度、准确度、流利度等方面更甚一筹。最后，多次任务练习有利于提高学习者的词汇表现。

本文基于心理语言学的注意和注意力的概念、Levelt 的话语模式、社会语言文化理论及其他相关的二语习得理论，对上述实验结果进行了理论层面的讨论，

并进一步提出了相关的任务型教学的建议。

本研究提出，在任务型语言教学实践和研究中，我们应当更多关注任务后阶段及其在聚焦于语言形式方面的作用。本研究建议 1）任务后书面抄录可被视作一种有效的任务后语言教学活动并应用于二语教学中；2）不同的任务后书面抄录条件对任务表现的影响各不相同；3）不同的任务类型对学习者的任务表现有不同的影响；4）多次任务练习有利于提高学习者的词汇表现。最后，本研究呼吁在以后的任务后研究中采纳“过程－结果”并重的实验方法以获得更加全面的实验结果及发现。

Acknowledgement

I am extraordinarily fortunate in having Dr. Peter Skehan as my Ph.D. supervisor. His expertise, understanding, and approach to life added considerably to my graduate experience. It was under his tutelage that I developed a focus concerning task-based research and became involved in the research of second language acquisition and psycholinguistics. His advice, supervision and crucial contribution made him a backbone of this research and so to this thesis. I am indebted to him more than he knows. Moreover, he is also a decent tennis player, cheerful hiker, experienced chef and overall a pretty good friend. The days together with Peter will definitely be the most valuable memory in my life. I owe him my eternal gratitude.

I would also like to convey my gratitude to the other members of my thesis committee, Dr. Jette G. Hansen Edwards and Dr. Gerald Nelson for the assistance they provided at all levels of the research. I gratefully acknowledge Dr. Pauline Foster from the St. Mary's University College for agreeing to be my external examiner despite her extremely busy schedule and for giving valuable feedback on my thesis.

Many thanks go to all my friends at the Chinese University of Hong Kong (CUHK) who helped me get through the three years of graduate school. Some friends, like Grace Leung, Amy Kong and Laura Man kindly introduced me the colourful local culture of Hong Kong. Some friends, like Jan Wang, Gavin Bei, Amy Dai and Bruce Jiang were always willing to share their bright thoughts with me, which were very fruitful for shaping up my ideas and research. Their jokes over lunch time and their excellent performance in Peter's research group made the life in CU wonderful.

I would also like to extend my thanks to English Department of CUHK for giving me the opportunity to study here. It is a pleasure to express my gratitude wholeheartedly to Tracy Liang for her support over the three years.

My parents deserve special mention for their profound love and consistent support. Their perseverance, hardworking and sincere love supported me to go through every kind of difficulties. I am especially grateful to my mother for nurturing my daughter with extraordinary caring throughout the long period of my study. Thanks also go to my young brother, for always being supportive and considerate.

Words alone cannot express my appreciation to my husband whose dedication, love and persistent confidence in me, has taken the load off my shoulders, and to my daughter for putting up with an absent mother and still being a happy and considerate friend of mine.

Finally, I would like to thank the participants who were important to the completion of the research, as well as expressing my apology that I could not mention personally one by one.

To my parents

-- I love you

Contents

List of Figures

List of Tables

Chapter One Introduction

1.1 Background of the study

In second language pedagogy, one of the major issues is what Stern (1983) called the 'code-communication dilemma' (Ellis, 2008). This is reflected by the dichotomy between "instructed" and "naturalistic" L2 learning. There are advocates of grammar teaching who view grammar instruction as the foundation for second language acquisition (Bialystock, 1990; Ellis, 1991; McLaughlin, 1990; Rutherford, 1988). By contrast, there are also advocates of the "zero option" (Ellis, 2008) which proposes to abandon formal instruction and to allow learners to construct their interlanguage naturally "through communication" rather than "for communication" (Krashen, 1982, 1985; Prabhu, 1987).

In the past two decades or so, researchers have come to realize that adopting a one-sided approach, either communicative-based or grammar-based, leads us nowhere. Theoretical paradigms on language teaching have changed accordingly. In addition to the two extreme positions, an alternative one has emerged which views formal instruction as facilitative for language acquisition. It is claimed that pure exposure to communication and focus on language use alone are not adequate for promoting a balanced L2 competence, and some kind of formal instruction needs to be incorporated into communicative contexts (Long, 1985; Long & Crookes, 1992; Long & Robinson, 1998). Some researchers argue that formal instruction does not necessarily enable learners to fully acquire what is

taught, but paves the way for its subsequent acquisition (Ellis, 2008) and "triggers the initial stages in what eventually results in grammar restructuring" (Gass, 1991:137).

Among the variety of proposals regarding the incorporation of formal instruction in communicative settings (see a review in Norris & Ortega, 2000), Focus on Form (FonF) has received increasing interest in the last two decades. A plethora of studies have been conducted either in L2 classrooms or in laboratory settings (Doughty, 2003; Doughty and Williams, 1998c; Ellis, 2001; Fotos & Nassaji, 2007; Norris & Ortega, 2000; Spada, 1997). There is now substantial evidence that when learners' attention is directed at linguistic forms and the meanings they encode—in the context of meaning-focused activities—learning takes place (see a review in Ellis, 2001, 2008). Focus on form is claimed to be necessary to push learners beyond communicatively effective language use toward target-like second language ability (Doughty & Williams, 1998b). It is, therefore, obvious that the focus on form option should be embedded in a communicative context to achieve both communicative competence and language ability. As a strong version of communicative language teaching, task-based instruction provides such a desirable setting for the operationalization of focus on form.

1.2 Brief rationale of the study

Based on a general agreement on the importance of a focus on form in a

communicative setting, researchers have explored how to achieve form focus in a variety of effective ways. Within a task-based framework (Ellis, 2003; Skehan, 1998), most of the previous focus on form studies concentrated on the pre-task or during-task stages. The focus on form options, such as pre-task planning, negotiation of meaning, feedback, and collaborative dialogue are operated before or during task performance (Ellis, 2003, 2008; Leeser, 2004; Lightbown & Spada, 1990; Loewen, 2005; Pica, 2002). Even so, there is no consensus concerning the effects of focus on form at pre or during-task stages. In contrast, surprisingly little attention has been given to the focus on form option at the post-task stage. In fact, there are good reasons for research into this pedagogical stage. Given the limited attention available for learners to attend to form during on-line communication (VanPatten, 1990), it is challenging to direct their attention to formal features prior to or during task performance. As Skehan (2007) suggests, at the post-task stage, the communication pressure is lessened. The meaning and the basic language needed to encode meanings has been established, so learners are more likely to switch their attention to the formal aspects of task performance. Further, the earlier performance provides the foundation for learners to reflect on their language system in a meaningful context.

To date, there have only been a few studies concerning focus on form at the post-task stage (Skehan & Foster, 1997; Foster & Skehan, forthcoming; Lynch, 2001, 2007), and most of those studies adopted post-task transcribing as the focus

on form activity. Given the relative infancy of the research, it is not surprising that there are some limitations in those studies. For instance, in Lynch's two studies, the sample sizes were not large enough for statistical analyses. Therefore, no statistical results were reported which caused problems in generalizing to other classrooms. In Foster & Skehan's (forthcoming) study, post-task transcribing was performed out of the class. It was possible that some unexpected intervening variables occurred when transcribing did not take place in front of the teachers. Also, in the previous studies, participants were engaged in post-task transcribing once or twice which may not be sufficient to demonstrate the treatment effects.

The above brief rationale highlights the need for further studies on focus on form at the post-task stage. In view of the significance of the post-task focus on form research and the scarcity of the existing research studies, the present research draws on and derives support form the previous related studies. This study mainly aims a) to explore the effects of post-task transcribing among a large number of participants (about 100 participants) in a relatively longer period (four sessions of post-task involvement); b) to investigate the impacts of different post-task transcribing conditions. Further, as a task-based research study, the present study also attempts a) to examine the different task effects between two distinctive task types (narrative vs. interactive tasks); b) to examine the role of multiple task practice (four sessions of task practice).

This research into how post-task involvement impacts on second language

improvement will cover many of the important issues in SLA—for example, the role of noticing, the process of speaking, reprocessing and modifying, the effect of interaction and pair work. In brief, it will afford insights into a series of important theoretical issues.

In addition, this study is intended not merely to shed light on how L2 improvement takes place as a result of post-task involvement and task practice. It is also motivated by a desire to discover what post-task conditions are most likely to facilitate language development. In other words, it has a pedagogical purpose. L2 teachers in a task-based context will find suggestions for specific procedures for the adoption of post-task transcribing in regular classrooms.

1.3 Organization of the Dissertation

This study consists of eight chapters. Chapter 1, the present chapter, relates the background and the brief rationale of the study, including the major controversies in second language pedagogy, the proposal for focus on form research and the neglect of focus on form research at the post-task stage. After the research gap has been identified, the multifaceted aims of the present research are presented.

Chapter 2 reviews literature relevant to this study. It reviews the focus on form research, including the context for its proposal, its definitions and features. The chapter then presents an overview concerning the features of task-based

research, the disadvantage of pure task-based instruction and the necessity of the integration of focus on form options into task-based instruction. It then reviews the focus on form research at pre- and during task stages. After the review of the theoretical background fc · post-task focus on form research, the chapter analyzes the previous post-task studies in great detail. It then outlines the motivation of the present study and the six research questions. This chapter also covers discussion on the fundamental issue for this study—the measurement of task performance.

Chapter 3 begins by describing the pilot studies undertaken prior to the main study. The chapter then describes the procedures for the main study, the participants in the study, the experimental tasks adopted, the research design, the setting of the experiments, the experimental procedures followed, the procedures for data transcribing and coding, and finally the methods of data analysis.

Chapters 4, 5 are concerned with data analyses and present the findings in terms of Research Questions one to four concerning the post-task transcribing effects. In Chapter 4, the narrative task performances are analysed in terms of fluency, accuracy, complexity and lexical performance, using multivariate analysis of variance (MANOVA). In Chapter 5 the interactive task performances are analysed similarly.

Chapter 6 analyses the task performance to address research questions five and six regarding the effects of multiple task practices and of different task types. Multivariate repeated measures are adopted to analyse the treatment post-task

groups and the control group separately.

Chapter 7 accounts for the results in the light of both cognitive psycholinguistics and the sociocultural theory. It pays attention to the nature of post-task transcribing, the role of interaction in the pair condition, the role of individual work at the post-task stage, the underlying process of task practice, and also the reasons for the distinctive task effects. It ends with a synthesis of both the findings of this research and offers the brief interpretation summary for those findings.

Chapter 8 presents both the theoretical and the practical significance of the present study. This chapter then provides several pedagogical implications which could be adopted by classroom teachers. Based on the analysis of the limitations, a series of suggestions are proposed for further studies. The chapter ultimately draws conclusions regarding the importance of focus on form research at the post-task stage.

Chapter Two Literature Review

2.0 Introduction

This chapter presents a review of the related literature. It begins with a review of the research on focus on form (FonF), including the context for its proposal, the features of FonF, and more importantly, the integration of FonF in task-based language teaching at pre-task, during task and post-task stages. Further, it discusses the theoretical background for the operation of focus on form activity at the post-task stage. Based on the critical review of the previous post-task studies, it outlines the motivation of the present study and the research questions that guide this research. Moreover, drawing on the literature of task performance measures in SLA, it discusses the measures of different aspects of task performance used in the present study.

2.1 Focus on Form in second language acquisition

2.1.1 Focus on Form: its background

In language teaching and second language acquisition (SLA) theory, the issue of formal instruction is at the heart of debate and has been subject to controversy and discussion among researchers for at least 40 years (Ellis, 2001). In the literature, there are two extreme options in language teaching: traditional grammar-based instruction and meaning-focused communicative language teaching. In the former, such as the grammar-translation method, the audiolingual

method, total physical response (TPR) etc, the second language is broken down into discrete items which are taught separately and step by step with the assumption that acquisition is a process of gradual accumulation of parts until the whole structure of language has been built up (Wilkins, 1976). The role of learners is to synthesize these parts into a whole unit and then use it in communication. This is what Wilkins (1976) termed as *synthetic* approach. Influenced by behaviorism and structural linguistics in 1950s and 1960s, the synthetic approach believes that "practice makes perfect", that is, language learning draws on pattern drilling and habit formation. Moreover, the language input is devised by teachers, instead of being derived from natural language use. The general principle for the traditional grammar-based instruction is that grammar is the foundation upon which language learning should be based (Hinkel & Fotos, 2002).

Research on both SLA and psycholinguistics shows that second language acquisition is not a process of accumulating entities (Long, 1991; Rutherford, 1988; Skehan, 1998; VanPatten, 1994). Language learning rarely happens with bits of language being learnt separately in an additive fashion. In addition, teachers can not predict and determine what students are going to learn at any given stage (Willis, 1996). Traditional grammar-based instruction ignored the language learning developmental processes through which L2 learners normally pass, and the fact that progress is not necessarily unidirectional (Ellis, 1994; 2008;

Long & Robinson, 1998). Furthermore, as for the authenticity of the input, Widdowson (1989) states that as fabricated by teachers, the classroom interactions and the practiced language forms will not necessarily transfer to actual language use in real-life situations. As a result, traditional grammar teaching has failed to prepare learners for spontaneous, contextualized language use. "The belief that a precise focus on a particular form leads to learning and automatization no longer carries much credibility in linguistics or psychology" (Skehan, 1996b, p.18).

Recognizing that treating the language purely as an object of study cannot develop the expected level of interlanguage proficiency, some researchers attempted to abandon grammar-based instruction in favor of more communicative-oriented language teaching which focused on language use (Nassaji & Fotos, 2004). They argued that formal language lessons would develop only declarative knowledge of grammar structure, not the procedural ability to use forms correctly, and that there was no interface between these two types of knowledge since they existed as different systems in the brain (see review in Dekeyser, 1998, 2001; Ellis, 2001, 2002). As such, they see formal instruction as unnecessary for interlanguage development.

Theoretically, this reflects Krashen's (1982, 1985) distinction between conscious learning and unconscious acquisition of language which involves entirely separate processes. Language acquisition is an implicit, subconscious process which occurs as a result of engaging in natural communication where the

focus is on meaning. By contrast, language learning is an explicit, conscious process which derives from formal instruction where the primary focus is on grammar and practice (Krashen, 1985). Krashen claims that learned or explicit knowledge which results from language learning cannot turn into acquired or implicit knowledge. According to Krashen, most of L2 cannot be taught; it must be acquired.

Proponents of communicative teaching claim that people of all ages learn languages best by experiencing them as a medium of communication. The essence of communicative language teaching is the engagement of learners in communication to allow them to develop their communicative competence (Long & Robinson, 1998; Widdowson, 1989; Savignon, 2005). Another tenet of communicative teaching is that exposing learners to large quantities of positive input that is comprehensible and meaningful is sufficient for language acquisition to occur. Grammar is acquired implicitly or incidentally (Reber, 1989, 1993; Krashen, 1985). This communicative language teaching approach underlies a variety of L2 classrooms, including those implementing Prabhu's procedural syllabus, Krashen's Natural approach, some content-based ELS instruction (e.g. immersion education), and task-based instruction.

However, research on the variations within communicative language teaching reveals at least the following problems (Long & Robinson, 1998): first, learning an L2 through experiencing its use is possible, but it is *inefficient.*

Learners who receive formal instruction of various kinds show higher levels of L2 proficiency than those only use the language (Doughty, 2003; Ellis, 1994, 2001, 2002, 2008; Lightbown, 2000; Long, 1983, 1988; Norris & Ortega, 2000; Spada, 2005). Put differently, the focus of communicative teaching on language use may have a ceiling effect on the acquisition of grammar; secondly, due to the maturational constraints on language learning and the adoption of communicative strategies instead of taking risks for more advanced language in communication, adult learners may become fluent, but not nativelike speakers, despite plenty of learning opportunity. Even worse, a pure communicative language teaching may lead to fossilization of L2 acquisition (Schmidt, 1993; Skehan, 1998).

Thirdly, things are not clearly better among child learners. Here it is worth mentioning the immersion programs initiated in Canada. Immersion programs provides students with a rich source of comprehensible input in which the teaching of a second language is integrated with the teaching of content subjects. The goals of immersion education include both academic achievement in the content subjects, and a high level of proficiency in the second language (Swain, 1985; Harley, 1989; Swain & Lapkin, 1995). However, even for the children who started the immersion education program in kindergarten, after many years of immersion, their productive skills remain far from nativelike, particularly with respect to grammatical competence (Swain, 1985; 1993; Swain & Lapkin, 1995), although their comprehension abilities are indistinguishable from those of native

speakers. The explanations commonly offered for this are mainly twofold: first, the immersion classroom context does not afford sufficient opportunities for "pushed output" (Swain, 1985). For example, students' output has been shown to consist of single words or short phrases with little opportunity for extended discourse, and feedback from the teachers on errors the students made seemed haphazard and random, based more on an 'irritation factor' than any rational pedagogical approach (Swain, 1988; Swain & Lapkin, 1995). Second, the immersion students are so focused on message content that they do not attend to form and consequently fail to acquire grammatical features that lack saliency (Ellis, 2003).

As such, the importance of attending to form becomes clear to both SLA researchers and teachers. Given that communicative language teaching by itself has been found to be inadequate (Ellis, 1997, 2002; Nassaji & Fotos, 2004), pedagogical interventions need to be interwoven into primarily communicative activities so as to overcome the limitations of both traditional grammar instruction and communicative language teaching (Doughty & Williams, 1998).

Researchers find it is necessary to look for an alternative approach, rather than to foster a single-sided teaching approach to promote both linguistic and communicative competence. In this context, motivated pedagogically by the findings of immersion and naturalistic research mentioned above and theoretically by the notion of noticing and the Interaction Hypothesis (Long, 1983, 1996),

focus on form (FonF) was proposed in an attempt to capture the strengths of the meaning-focus communicative approach while dealing with its limitations (Long & Robinson, 1998). In the Interaction Hypothesis, interaction between learners and other speakers, especially more proficient speakers, is of crucial importance for language development. Negotiation of meaning occurring in interaction produces negative feedback (e.g. recasts, a corrective reformulation of learners utterances) to draw learners' attention to mismatches between input and output, thus induce them to notice the forms which are not only comprehensible but also meaningful (Long & Robinson, 1998). As such, negative feedback during negotiation of meaning may facilitate L2 development (Long, 1996). Motivated by the role of negotiation of meaning during interaction, Long (1991) proposed the option of Focus on Form to be incorporated in meaning-based communicative language teaching as an alternative to either traditional grammar instruction or pure communicative language teaching.

Since its conception, the idea of FonF has been widely advocated in the SLA literature. However, due to its popularity among researchers and teachers, there is considerable variation in how the term 'focus on form' is understood and used, and the construct has been interpreted and used differently by different researchers and teachers. At the outset, it is necessary to clarify the construct of FonF prior to its application to the present study.

2.1.2 Focus on Form: its definition and features

In his seminal work, Long (1991) initially introduced the notion of focus on form: 'focus on form...overtly draws students' attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning or communication' (Long, 1991, 45-46). Based on this theoretical notion, Long & Robinson (1998) later raised a more pedagogically applicable definition of Focus on form as: 'focus on form consists of an occasional shift in attention to linguistic code features—by the teacher and/or one or more students--triggered by perceived problems with comprehension or production' (Long & Robinson, 1998, p23).

The notions include minimally the following central features of FonF (Doughty & Williams, 1998c; Williams, 2005):

- A primary emphasis on the communication of meaning;
- A brief attention shift from that emphasis on communication to focus on language as object;
- A problem-oriented trigger for the attention shift

The idea of attention to form differs from explicit formal instruction. The definitions and the central features of FonF make focus on form distinguishable from the notion of focus on forms (a term for traditional grammar teaching in Long, 1991) and focus on meaning (a term for meaning-focused communicative language teaching in Long 1991). Focus on form *entails* a focus on formal

elements of language with the prerequisite engagement in meaning, whereas focus on formS is *limited* to such an isolated language focus with no intention for effective communication, and focus on meaning *excludes* much effort concerning language elements (Doughty & Williams, 1998b, p. 4).

In Long's definition, FonF would be viewed as a responsive teaching intervention for online occasional attention shifts to problematic form in a communication setting. The problematic form occurs incidentally as a breakdown or error of some difficulty in either production or comprehension. From a pedagogical perspective, this means that teachers should wait for issues to emerge and respond to them as needed.

This original conception of FonF has been broadened to the extent of allowing for a planned, and even separated instructional focus on form that need not be derived from a real-time problem-trigger (Doughty & Williams, 1998). In Doughty and Williams' definition, FonF includes both reactive and proactive FonF. Reactive FonF is consistent with Long's original definition. A reactive FonF would require that the teacher notice and be prepared to handle various learning difficulties as they arise. Reactive FonF would seem to be most congruent with the general aims of communicative language teaching. However, in the majority of language class settings, a more proactive FonF approach is likely to be feasible as well (Doughty & Williams, 1998c). In proactive FonF, the teacher must make a determination in advance as to which form— among the language forms that are

potentially good candidates for focus—to select for attention at any particular time. Based on considerations of individual differences, developmental language learning sequences, input quality, formal and functional complexity, and L1 influence on SLA processes, the suggestion for proactive FonF should be distinguished from that of the traditional structural approaches in which isolated language forms are selected and sequenced on the basis of intuition (Doughty & Williams, 1998c). Proactive FonF emphasizes the design of tasks that ensure that opportunities to use problematic forms while communicating a message will indeed arise (Doughty & Williams, 1998c, p. 211).

In addition to the reactive-proactive distinction, Ellis, Basturkmen, and Loewen (2001) use the term *preemptive* FonF. In this case, the trigger for the attentional shift from meaning to form may not be an actual or perceived problem; instead, it may simply be a problem that is anticipated. In this way, no problem-trigger is required. However, some researchers question whether it is a FonF at all (Williams, 2005). In Ellis et al's (2001) teacher-initiated preemptive FonF, this focus is purely incidental and the teacher decides on the spur of the moment that learners might encounter difficulty processing a word or form, and therefore he/she launches into a brief, explicit instructional sequence. The 'problem' is only anticipated, and unfortunately as Ellis et al show, the prediction of the problem is not always accurate. It therefore cannot be treated as a real problem trigger. However, as we mentioned above, the problem-trigger (or the

problematicity) is one of the three central features of FonF. This is an example to show that the term FonF is used and interpreted differently to a large extent to meet the different requirements of various studies, thus some misinterpretation of or confusions with the notion are worrying for current FonF research.

Focus on form has received sufficient attention in the L2 teaching field as well. Willis (1996) and Willis & Willis (2007) propose a clear-cut distinction among focus on meaning, focus on language and focus on form. In their opinion, a focus on language is the option in which "learners pause in the course of a meaning-focused activity to think for themselves how best to express what they want to say, or a teacher takes part in the interaction and acts as a facilitator by rephrasing or clarifying learner language" (p.5). In contrast, a focus on form is an option in which one or more lexical or grammatical forms are isolated and specified for study, or in which the teacher comments on student language by drawing attention to problems. Similar to Long's definition, a focus on meaning is an option in which participants are concerned with communication (Willis & Willis, 2007, p.5).

Willis & Willis (2007) viewed their definition of focus on language as similar to that of focus on form by Long (1991, Long & Robinson, 1998), and their notion of focus on form as similar to that of *focus on formS*. However, a careful scrutiny on the context of Willis and Willis' definitions reveals that their notions of both focus on language and focus on form seem to be within Long's

notion of focus on form. Doubtless focus on language is a FonF option in that it meets all the central features of FonF: it is launched during communication where the meaning is the center of the activity; the focus is shifted to language at some time which is triggered by the breakdown of the communication.

Let us turn our attention to the notion of focus on form which Willis & Willis (2007) treated as similar to Long's focus on formS. The question is whether it is similar to Long's focus on formS or focus on form. In Willis & Willis' focus on form, "teachers direct learners' attention to specific forms which occur in the course of a task or an associated text. … Teachers should take care that this focus on form does not detract from a focus on meaning" (p.5). All the language forms in the focus are taken from texts that the learners have processed in the course of the task sequence. This means the forms they focus on are rooted in a meaning context. This will not only ensure that the meaning of these forms is more easily recognized, it also makes them more memorable. And for the problematic forms, teachers use corrections, and provide negative feedback as a part of a form-focused activity (Willis & Willis, 2007). In brief, although Willis & Willis' focus on form is a separate pedagogical intervention rather than a brief attentional diversion, its aim is to keep the primary focus on meaning and the use of focus on form is "triggered by an analysis of learner need rather than being imposed externally by a linguistic syllabus" (Doughty& Williams, 1998b, p.8). It is therefore acceptable to view this focus on form as compatible with Long's focus

on form rather than focus on formS in which the language focus are selected depending on the analyst's linguistic preference and sequenced in linear and additive fashion according to such criteria as (usually intuitively assessed) frequency, valence or difficulty (Long & Robinson, 1998).

The varied interpretations and plethora of rather similar terms concerning focus on form have broadened our understanding of the notion of FonF (see review in Doughty & Williams, 1998; Norris & Ortega, 2000). As Doughty & Williams proposed, not limited to Long's implicit online FonF option, the pedagogical operation of FonF can be more flexible and diverse. It is reasonable to view FonF flexibly as a continuum from a most implicit, unobtrusive pole to an explicit obtrusive pole while meaning is always at the center of the activity and a problem trigger does exist (Doughty & Williams, 1998c). It is furthermore entirely possible to combine explicit and implicit FonF techniques, depending upon the particular acquisition circumstances (Doughty & Williams, 1998c). There is growing evidence of the effectiveness of such a flexible curricular approach involving a variety of successful task-technique combinations (see review in Norris & Ortega, 2000).

In the present study, the notion of focus on form is adopted by integrating the notions of FonF from both SLA research (Doughty & Williams, 1998; Long, 1991; Long & Robinson, 1998) and pedagogy (Willis, 1996; Willis & Willis, 2007). In particular, focus on form is viewed as a reactive option operating at the post-task

phase in task-based language teaching.

In the following sections, in the first place the literature which is concerned with focus on form in task-based language teaching will be reviewed within a task implementation framework. And then based on a critical view of the previous studies, the rationale and the motivation of the present study in terms of both theoretical and pedagogical aspects will be outlined.

2.2 Focus on Form in Task-based Language Teaching

So far, we have reviewed the background to proposals for a focus on form, its definitions and features. Unlike the extreme teaching approaches which stress either full formal instruction or pure communication within the L2, focus on form is an integration of both of these approaches. In other words, focus on form cannot exist in a vacuum, but should be embedded in a meaning-focused communicative setting. From a communicative perspective, the most effective way to assist language learning in the classroom is through communicative tasks (Nassaji, 2000). That is to say, tasks are an important component of communicative language teaching (CLT) (Ellis, 2003). Task-based language teaching, as a strong version of CLT, provides an optimal setting for the incorporation of focus on form (Doughty & Williams, 1998; Long, 1991).

2.2.1 Task-based language teaching

Task-based language teaching refers to an approach based on the use of tasks as the core unit of planning and of syllabus design in language teaching (Ellis, 2003; Long & Crookes, 1992). The last twenty years have seen an enormous growth of interest in task-based language learning and teaching (Bygate, Skehan & Swain, 2000a; Ellis, 2003; Samuda & Bygate, 2008; Skehan, 1998; Willis, 1996; Willis & Willis, 2007). This interest has been motivated to a considerable extent by the fact that 'task' is seen as a construct of equal importance to second language acquisition (SLA) researchers and to language teachers (Pica, 1997). For researchers, tasks can be used as basic conceptual units to analyse learning behaviours that lead to second language acquisition (Bygate, 1996, 2000; Dornyei & Kormos, 2000). In particular, tasks "allow researchers to break down the complex, prolonged learning process into discrete segments with well defined boundaries, thereby creating meaningful 'anchor points' in discussing the various dimensions (e.g. cognitive, affective or socio-dynamic) of L2 processing" (Dornyei & Kormos, 2000, p 276). For language teachers, tasks can be the device for organizing the content and methodology of language teaching (Prabhu, 1987). Furthermore, in communicative language teaching, tasks can function as a useful device for planning a communicative curriculum, particularly in contexts where there are few opportunities for more authentic communicative experiences, for example many FL situations (Ellis, 2003), and tasks themselves serve as

communicative language activities.

Then what exactly is a task? In the literature, various definitions have been offered that differ quite widely in scope and formulation (Van den Branden, 2007). Bygate, Skehan & Swain (2001) pointed that "definitions of task will need to differ according to the purposes for which tasks are used". They proposed a 'basic, all-purpose definition' which can be modified to reflect the different purposes of tasks—that is "a task is an activity which required learners to use language, with emphasis on meaning, to attain an objective" (p.11).

Among various specific definitions, Skehan (1998) and Nunan (1993)'s definitions are conveniently broad and can be recognized by both L2 researchers and teachers.

According to Nunan (1993), *A task is ' a piece of classroom work which involves learners in comprehending, manipulating, producing or integrating in the target language while their attention is principally focused on meaning rather than form* (Nunan, 1993, p.59).

In Skehan's definition, a task is an activity in which

-- meaning is primary;

-- there is some communication problem to solve;

-- there is some sort of relationship to comparable real-world activities;

-- task completion has some priority;

-- the assessment of the task is in terms of outcome (Skehan, 1998, p.95)

One of the key features which are shared by all the above notions of task is the primary focus on communicative meaning. A 'task' requires the participants to function primarily as language users in the sense that they must employ the same kinds of communicative processes as those involved in real-world activities (Ellis, 2003). The priority of communicative meaning makes task-based instruction appealing in that the authentic classroom communication is likely to lead to a desirable communicative success in real world situations.

However, that does not mean task-based instruction is perfect. Some researchers (Skehan, 1996a; Skehan & Foster, 2001) argued that such an emphasis on meaning can unfortunately result in a series of problems for learners' interlanguage development. First, putting great emphasis on communicating meaning, learners may heavily rely on some elliptical or incomplete sentences (which is frequent in native speakers communication) and do not care about the correctness or completeness of what is said (Skehan, 1996a). Second, in respect of language comprehension, learners may naturally and inevitably adopt comprehension strategies in their attempt to understand the intended meaning in the message. In this way only partial forms as a clue to meaning are needed to be processed. However, "processing language to extract meaning does not guarantee automatic sensitivity to form, and the consequent pressures for interlanguage development..." (Skehan, 1996a, p41). Furthermore, with regard to language production, communicative strategies are the preference for learners when faced

with communicative pressures. The use of cognitive and linguistic communicative strategies can make learners lessen the cognitive load, handle communicative pressures, and comfortably get meaning across in the communication, but in so doing learners 'escape' from the required language engagement which is necessary for interlanguage development (Skehan, 1996a). Even worse, the effective use of communicative strategies at the onset of learning may become proceduralized and reused on other occasions (Skehan, 1996a). This kind of proceduralization may encourage a comfortable fossilization rather than to promote interlanguage development (Skehan & Foster, 2001).

2.2.2 Focus on Form at pre-task and during task stages

Based on the foregoing discussion of the defects of task-based instruction, Skehan (1996a, Skehan & Foster, 2001) suggest that

> it may not be possible to rely on a task-based approach to automatically drive interlanguage forward...it is necessary, if task-based approaches to instruction are to be viable, to devise methods of focus on form without losing the values of tasks as realistic communicative motivators, and as opportunities to trigger acquisitional processes (Skehan, 1996a, p42).

This is a most important motivation for the incorporation of focus on form in task-based instruction. In recent decades, in task-based research and pedagogy, arguments have been made to interweave a concern for language form without compromising the communicative nature of the task (Bygate, Skehan & Swain, 2001; Doughty & Williams, 1998; Ellis, 2001, 2002, 2003, 2005; Fotos & Nassaji,

2007; Garcia Mayo, 2007; Norris & Ortega, 2000; Skehan, 1996a, 1998, 2003; Skehan & Foster, 2001; Willis, 1996; Willis & Willis, 2007). To add an option of FonF in task-based language teaching is appealing for the implementing of FonF in a communicative language teaching context (Skehan, 2007).

Researchers vary in their proposals for how this focus on form in task-based instruction can be best achieved. Within a task implementation framework (Ellis, 2003; Skehan, 1998), much of the task-based focus on form research falls into three categories: focus on form at pre-task, during task and post-task stages.

With regard to pre-task research, as Skehan (2007) pointed out, most task-based researchers, instead of investigating a wide range of pre-task activities (Willis, 1996), emphasize one influence only: pre-task planning (Crookes, 1989; Ellis, 1987;Foster & Skehan, 1996, 1999; Mehnert, 1998; Mochizuki & Ortega, 2008; Ortega, 1999; Sanguram, 2005; Skehan & Foster, 1997, 1999; Tavakoli & Skehan, 2005;Wigglesworth, 1997; Yuan & Ellis, 2003). As far as the pre-task planning effect on language form (i.e. complexity and accuracy) is concerned, the research shows that pre-task planning has a stronger effect on fluency and complexity than on accuracy which means that during pre-task planning, learners give more attention to the content of the following performance rather than to the detailed linguistic forms used in the speech (Ellis, 2005).

With respect to focus on form at the during-task stage, researchers are mainly concerned with the effect of the interaction process, such as the efficacy of

negotiation of meaning (Mackey, 1999; Pica, 1994, 1997; Pica et al, 1989; Polio & Gass, 1998; Van den Branden, 1997), and of corrective feedback, such as recasts, explicit corrective feedback etc (Ammar & Spada, 2006; Doughty & Varela, 1998; Lyster, 2004; Lyster & Ranta, 1997; Mackey & Philp, 1998; Truscott, 1999). The research on negotiation of meaning shows that it is beneficial for learners' comprehension of input and then for their language development. In a recent meta-analysis on the effectiveness of corrective feedback(CF) in terms of grammar, Russell & Spada (2006) find support for the effectiveness of corrective feedback for L2 grammar learning. Further, to some extent the benefits of CF are durable to the delayed post-tests. However, some researchers are doubtful about the effects of negotiation of meaning and recasts. They claim that it is likely that benefits of negotiation, whatever they might be, are not uniform (Pica, 2002). As such, it can not be treated as a particularly dependable pedagogical technique. Pica (1997, 2002) and others (Foster, 1998; Musumeci, 1996; Williams, 1999) have claimed that negotiation in the classroom, either among learners or between learners and the teacher, is not as frequent as might be wished. Nor can transfer-of-information tasks guarantee that learners will negotiate (Foster, 1998). As for the effects of corrective feedback, Truscott (1996, 1999, 2007) argues that the error correction in oral performance does not improve learners' ability in terms of grammar.

In brief, although there are heated debates concerning the effectiveness of

pre-task and during task activities as ways of promoting focus on form, the research has definitely deepened our understanding concerning the incorporation of focus on form in task-based language teaching. However, the understanding cannot be complete at all without the research on another indispensable phase of task-based instruction: the post-task stage. The following sections will concentrate on this stage which has received the least attention from task researchers.

2.2.3 Focus on form at post-task stage

As compared to the numerous studies concerning focus on form at the pre-task or the within task stages, focus on form at the post-task stage has been surprisingly neglected by most task researchers to date (except Forster & Skehan, forthcoming; Lynch, 2001, 2007; Skehan & Foster, 1997). However, Doughty (2001) has argued that the "most compelling proposal for a solution to the timing issue (of focus on form) involves tapping the powerful cognitive resources that enable learners to make use of a *recently occurring* utterance" (emphasis added, p.252). At the post-task stage, activities would generally have the function of 'highlighting form in the earlier task performance' to bring it more into focus (Skehan & Foster, 1997). The idea of focus on "earlier task performance" is compatible with some conversational analysis results (Schenkein, 1980, p.46) which reveal that "systematic use of *resources from prior talk* in current talk apparently organizes the conversation" (emphasis added). Skehan (2007) claims

that the rationale for the post-task phase is that "the previous active communicative activity has prepared the ground for learners to reflect upon what they have done, and engage in analysis, reorganization of their language system, and consolidation of the progress they have made" (p.63).

To give a more developed rationale concerning the above statement, the following section discusses the theoretical background for the post-task activities.

2.2.3.1 Theoretical background of Focus on Form at post-task stage

Focus on form concerns how the learner's attentional resources are allocated in a communication setting. In particular, focus on form attempts to shift students' attention to linguistic forms while their primary focus of attention is on meaning. Clearly, the notion of attention and noticing is of crucial importance for the understanding of focus on form at the post-task stage. As such, the conceptual rationale is reviewed first with the two related psychological constructs — attention and noticing (with Schmidt's Noticing Hypothesis). Furthermore, in previous related studies, focus on form at the post-task stage is concerned with the reproduction of language in different ways. Levelt's Speaking Model (Levelt, 1989), and Swain's Output Hypothesis (Swain, 1985, 1995, 2005) are reviewed to provide the basis for the employment of post-task activities which may promote effective focus on form.

Attention and Noticing. As far as attention is concerned, the classic view in psychology is that limited capacity is the primary characteristic of attention (Schmidt, 2001), and this view has been taken on by many researchers in SLA (Anderson, 1993; Mackey, 1999; McLaughlin, Rossman & McLeod, 1983; Gass, Svetics & Lemelin, 2003; Skehan, 1998;VanPatten, 1990).

Due to limited attentional resources and the competition between form and meaning in communication, attention must be strategically allocated (VanPatten, 1990, 1994, 1996). VanPatten has argued that what is important in most SLA contexts is communicative meaning. Limited attentional resources are directed first at those elements that carry message meaning, primarily lexical, and only later, when the processing demands for meaning come down, towards communicatively less informative formal features of language (VanPatten, 1990, 1994, 1996). In other words, attention is not only limited, but also selective (Schmidt, 2001). Learners can acquire forms only when processing for meaning is automatic and attention is freed and available.

In this vein, the basic assumption of focus on form is nicely compatible with the notion of limited and selective attention, as it is claimed that the fundamental assumption of focus on form is that "meaning and use must already be evident to the learner at the time that *attention* is drawn to the linguistic apparatus needed to get the meaning across" (Doughty & Williams, 1998, p. 4). In particular, the effects of focus on form resulting from post-task activities are likely to be

expected in this context. With limited and selective attention, it is clear that language learners are "not free to notice whatever they want whenever they want and that a number of factors influence noticing ability" (Schmidt, 1990). The post-task phase occurs as a desirable stage for noticing. After the completion of a task, learners' attention has been freed up from communicative pressure, and they have sufficient cognitive resources to attend to/notice the language forms which occurred in the previous task performance (Skehan, 1996b, 2007; Lynch, 2007).

Although noticing does not necessarily mean development or change (Skehan, 2007), Schmidt (1990, 1993, 2001) proposes that conscious attention to form or what he calls "noticing" is a necessary condition for language learning. Holding a strong version of the claim that attention is central, Schmidt (2001, p3) emphasized that "the concept of attention is necessary in order to understand virtually every aspect of second language acquisition". Most SLA researchers agree that noticing or attention to language forms plays an important role in L2 learning (Bygate, Skehan & Swain, 2001; DeKeyser, 1998; Doughty, 2001; Ellis, 2001, 2002; Robinson, 1995, 2001a; Skehan, 1998; Swain & Lapkin, 1995, 2001).

A subsequent question comes up: which forms are to be attended to? In Schimdt's Noticing Hypothesis, it is stated essentially that "learners need to pay attention to (and notice) details and differences in order to learn" (Schmidt, 2001). The processing of noticed details has been referred to as *noticing the gap* (Schmidt & Frota, 1986). The gap is the mismatch between the input and the

learners' own output (Schmidt & Frota, 1986). As Long (1996) pointed out, Schmidt's Noticing Hypothesis, emphasizing the importance of noticing the form in the input, is mainly input-oriented. Researchers have shown interest mostly on how to process input effectively so as to cause it to be noticed for interlanguage acquisition (VanPatten, 1996; Gass, 1997).

In the last decade, the noticing focus has been enlarged to be concerned with not only input but also output, and as a result the concept of noticing has been broadened to a wider range: *noticing the gap* and *noticing the hole* (Swain, 1995, 1998). These two notions appear to be related, but differ in crucial ways (Swain, 1995; Doughty & Williams, 1998; Williams, 2005; Izumi et al, 1999).

Noticing the gap occurs when learners notice that their interlanguage (IL) differs from the target language. The underlying assumption is that learners already have an established IL form and they figure out that the established IL form is an error. It suggests a simultaneous presence of both IL and TL forms. *Noticing the hole* takes place at the point at which learners realize that they do not have the means to say something that they want to say (Swain, 1995). This differs from *noticing the gap*, in that while noticing the hole, learners may not yet have developed an IL form to express what they want, that is there may be an absence of IL form (Izumi et al, 1999).

As such, noticing the gap and noticing the hole take place at different acquisitional stages and involve different cognitive processes. Activities that

promote noticing the hole seek to intervene at the point at which input becomes intake, an earlier stage in the acquisition process (Williams, 2005). In contrast, activities that facilitate noticing the gap attempt to destabilize the IL and move it toward more targetlike accuracy, a later stage in acquisition, with a combination of positive and negative evidence (Doughty, 2001). In the two acquisition stages, different cognitive processes are involved: noticing the gap requires cognitive comparison (Doughty, 2001). This means incoming input would have to be compared either to representations stored in long-term memory or to traces left in working memory (for the fine-grained analysis, see Doughty, 2001). In other words a comparison between an IL form and TL form is inevitable. Then, once the mismatches are noticed and compared, it may require a destabilization and restructuring of IL knowledge, and consequently the formation of new form-meaning connections. Noticing the hole on the other hand, would seem to be a simpler process in that "it does not require comparison to representations that have been previously stored" (Williams, 2005: 682). It mainly involves the recognition of a lack in IL and the attempt to establish such a new IL form, although there is no guarantee that the hole will be filled by a TL form (even though this is clearly a pedagogical goal) (Williams, 2005).

With regard to focus on form in task-based instruction, both noticing the gap and noticing the hole can be involved. The problem trigger for a focus on form may be an error which would launch noticing the gap or the situation when

learners encounter a meaning they do not know how to express which would promote noticing the hole (Izumi & Bigelow, 2000; Shehadeh, 2001; Swain & Lapkin, 1995; Swain, 1998; Williams, 1999). Is it possible for learners to be engaged in a certain situation where both noticing the gap and noticing the hole are involved? The post-task phase provides such a suitable platform for both noticing the gap and noticing the hole —for the former, learners have the chance to review their own previous task performance so as to identify and notice the language errors (or the difference from the target language) and to give some offline feedback on those errors; for the latter, learners may carry out some subsequent activities (e.g to redo the task in various ways) to analyze and reflect upon the inadequacy (or absence) in their interlanguage as they are no longer preoccupied with formulating meaning. In other words, they may notice the need for new forms, or the relevant forms in interlocutors' language (Lynch, 2007; Skehan, 2007). As such, the gaps and the holes which have been noticed become vital stimuli for future learning and interlanguage development.

Levelt's Speaking Model. Different from the pre-task activities (e.g. pre-task planning), most of the post-task activities involve the production of language in written (e.g. to transcribe the earlier performance in Clennell, 1999; Foster & Skehan, forthcoming; Lynch, 2001, 2007; Mennim, 2003;) or in oral forms (e.g. to redo the task publicly in Skehan and Foster, 1997). Those post-task activities are

proposed as a way of focusing learners' attention on language formal features for language development. Levelt's speaking model and Swain's Output Hypothesis provide grounds for the production of language at the post-task stage.

According to Levelt (1989), speech production consists of three major components: the Conceptualizer, the Formulator, and the Articulator. The Conceptualizer includes conceiving of an intention, selecting the relevant information, ordering this information for expression, keeping track of what was said before, and so on. The Conceptualizer formats communicative intentions in such a way that the formulator can handle them. The production of the Conceptualization stage is a preverbal message. The Formulator converts the preverbal message into a speech plan which involves two major processes that translate a conceptual structure into a linguistic one. The first step is the grammatical encoding of the message, a process consisting of procedures for accessing lemmas and of the syntactic building procedures. Levelt argues that a lemma is activated when its meaning matches part of the preverbal message and that this activation leads to certain syntactic building procedures. Along with the relevant lemmas, the syntactic building procedures produce a surface structure. The second step of formulating is phonological encoding. This includes retrieving or building detailed phonetic and articulatory planning for lemmas and for the whole utterance. The output of the Formulator is a phonological plan, priming the speaker for the articulation of the utterance. The phonological plan is executed in

the third major processing component, the Articulator. Articulating is the execution of the phonetic plan through retrieving chunks of internal speech from a temporary storage called the articulatory buffer and through motor execution. This is a component in which speakers' internal linguistic knowledge turns into audible sounds (for a detailed review, see Muranoi, 2007).

Levelt (1989) argues that conceptualizing involves highly controlled processing that demands attentional resources for both L1 and L2 speakers. For L1 speakers, other processing components such as formulating and articulating are claimed to be largely automatic, and demand very little executive control. However, for L2 speakers who have not yet developed full command of the target language, formulating and articulating are likely to be controlled. Second language production is bound to place strains on learners in all the three components (Muranoi, 2007): the Conceptualizer, the Formulator and the Articulator when there is time pressure. SLA researchers are concerned with how to provide learners with a learning experience which will enable them to manage the form-meaning relations, that is to enhance the efficiency of the formulating so as to find an appropriate form to match the preverbal message. Post-task activities may provide supportive conditions for such form-meaning management in the following two ways:

First, as in the case in task repetition (Bygate, 1996, 2001; Gass et al, 1999), since they have done the task before, when doing the post-task activity, learners

have been informed of the task topic and have already tried to segment the preverbal message for speech production. That means they are familiar with the conceptualization process for speaking, and so most of their attention can be concentrated on the selection and articulation of language. In the post-task condition, most attention is available for the formulating component so that the form-meaning connection is more likely to occur correctly and more appropriately than in the pre/during task condition.

Second, when redoing the task at the post-task stage, learners have already experienced the formulation and articulation of the speech. In other words, the post-task activity aims to lead them to go through the formulation and articulation processes a second or a third time. Drawing on the encodings which they had previously used, learners may benefit potentially in the following three ways (Bygate, 2001; Foster & Skehan, 1996; Skehan, 2001):

i) if they are to redo the task orally, they may become more fluent since the necessary lemmas and syntactic patterns are likely to be easier to be accessed in the second time;

ii) alternatively, if they are to transcribe their previous speech, they may not gain much in fluency, but might pay more attention to the precision of the previous speech by means of correcting the errors or refining the speech to match the norms of the target language.

iii) a third possibility is, in subsequent activity after task performance, the

learner might choose to build on the routines established previously to produce a more complex or more sophisticated conceptualization or formulation of the message.

In short, according to Levelt's speaking model, L2 speakers have constraints on all three components of L2 language production. The operation of post-task activities provides the opportunity for learners to reduce the pressure on all the three components of speaking: the conceptualizing, the formulating and the articulation, especially the latter two. Without much pressure in terms of attentional allocation between form and communicative meaning, the efficacy of the Formulator and the Articulator has been enhanced. As such, it may channel the learner's attention to formal features of the speech and give rise to a change in task performance.

Swain's Output Hypothesis. Levelt' s Speaking Model accounts for the possibility that a post-task activity may give rise to a change in L2 performance. Researchers are further concerned with the question: whether the changes or whether producing speech can lead to L2 acquisition and development. Swain's Output Hypothesis (1985, 1995, 1998, 2005) tries to account for how L2 output affects the various cognitive processes involved in SLA, and her Output Hypothesis is viewed as an important extension of theories that consider input as the most important aspect of second language acquisition (Swain, 1985).

Swain argues that output is not only a product of language acquisition outcome, but also a part of the *process* of learning, and therefore she suggests pushing learners "...toward the delivery of a message that is not only conveyed, but that is conveyed precisely, coherently, and appropriately" (Swain, 1985, p248-249). The Output Hypothesis proposes three functions for output (Swain, ibid):

(1) *The Noticing/Triggering Function* It is claimed that under some circumstances, the activity of producing the target language may prompt L2 learners to recognize consciously some of their linguistic problems. "It may bring their attention to something they need to discover about their target language (possibly directing their attention to relevant input)" (Swain, 2005, p474). This awareness triggers cognitive processes for learners to generate new linguistic knowledge or consolidate their existing knowledge (Swain & Lapkin, 1995). Unlike comprehension, production does not allow learners to rely on external cues and general nonlinguistic knowledge (De Bot, 1996). In production, learners will inevitably discover what they can and cannot do. Noticing a problem is not solving it, but the awareness of a problem may lead to more attention to relevant information in the input. Thus, noticing can lead to learning (De Bot, 1996).

(2) *The Hypothesis Testing Function* The claim is that from the learner's perspective, output may sometimes be a 'trial run' reflecting their hypothesis of how to say (or write) their intent (Swain, 2005, p.476). In other words, producing

output is potentially a way of testing one's own hypothesis about the target language. Research shows that learners modify their output in response to conversational feedback (Pica et al, 1989; Loewen, 2005). Swain (2005) argued that if learners were not testing hypotheses, then the modified output would not be expected following feedback. Further, it is assumed that the processes in which learners engage to modify their output in response to feedback are part of the second language learning process (Swain, 2005). In modifying their output, learners "test hypotheses about the second language, experiment with new structures and forms, and expand and exploit their interlanguage resources in creative ways" (Pica et al, 1989:64). Studies show that pushing learners to modify their output not only results in better control of features which had already been acquired (Nobuyoshi & Ellis, 1993); but also leads learners to be engaged in acquisition of new forms (Pica et al, 1989).

(3) *The Metalinguistic (Reflective) Function* It is claimed that "as learners reflect upon their own target language use, their output serves a metalinguistic function, enabling them to control and internalize linguistic knowledge" (Swain, 1995, p.126). In other words, output may cause the learner to engage in a more syntactic processing than is necessary for the comprehension of input. In their work, Swain and her colleagues (Swain, 1995; Swain & Lapkin, 1995) have developed various classroom activities termed as "collaborative dialogue" in which students have to work together to solve "form-related"

problems in the target language. Through collaborative discussion, the students become more aware of a language problem and try to solve it together. In this setting, the elicitation of relevant input seems to be the mechanism through which learning can take place (De Bot, 1996).

The essential difference between language production during the task and after the task is that the former mainly aims at getting the job done—to accomplish the task, while the latter is to redo the task in various ways for language reflection, consolidation and development (Lynch, 2007; Skehan, 2007; Wills, 1996). In the post-task condition, without much online communication pressure, learners may process the language more precisely, coherently and appropriately (Swain, 1995, 2005) which reflects the three functions of the output in the following ways:

First, at the post-task stage, there is an existing earlier performance to help learners to notice the gap or the hole. To redo the task orally or in written form provides opportunities for learners to notice the distance between their interlanguage and the target forms. In other words, the post-task output processing pushes learners further in their cognitive processing (e.g. the cognitive comparison between their own speech and the target norms) (Doughty, 2001) and prompts them to perceive or conceive the utilized structure.

Second, post-task activities may involve not only the production of language, but also modification of the previous speech (Lynch, 2001; 2007). As discussed

above, the process of modification, in essence, is a process of hypothesis testing for L2 language development. In the process of hypothesis testing in response to internal or external feedback, "if learners cannot work out a solution, they may turn to input, this time with more focused attention, searching for relevant input. Or they may work out a solution, resulting in new, reprocessed output. What goes on between the first output and the second is part of the process of second language learning (Swain & Lapkin, 1995). As such, it will be beneficial both for the consolidation of the existing knowledge of L2 and for acquisition of the new forms.

Third, focus on form at the post-task stage can be operationalized either individually or in pairs (Lynch, 2001, 2007). The pair work, termed "collaborative transcribing" in Lynch's studies (2007), is similar to the "collaborative dialogue" proposed by Swain and Lapkin (2001). Researchers find that when the activities are performed collaboratively, they lead to a focus on form as students engage in constructing the meaning required by the task (Kowal & Swain, 1997; Swain, 1998). In collaborative post-task activities, learners are likely to be involved in metalinguistic processing, and this processing may also lead to modified and reprocessed output— a possible step toward language acquisition (Izumi et al, 1999).

In sum, Levelt's Speaking Model and Swain's Output Hypotheses lend theoretical support for the research on focus on form at the post-task stage.

Post-task focus on form activities may not only give rise to changes in learners' performance, but, more importantly, are likely to lead to language acquisition by means of noticing, hypothesis testing and metalanguage reflection. Thus, it is promising to investigate focus on form at the post-task phase so as to explore its role on learners' performance improvement and furthermore, their interlanguage development.

Nevertheless, despite the above sound reasons for the research on focus on form at the post-task stage, the empirical studies in this regard are surprisingly few. In the next section, we will review previous studies concerning post-task focus on form. Those studies, although small in number, provide the practical grounds for the motivation of the present research.

2.2.3.2 Previous studies on focus on form at post-task stage

There is a paucity of research with regard to the post-task activity for focus on form effects. One of the reasons may be that in the research context, some researchers viewed task-based instruction as consisting only of pre-task activity and task performance (or only task performance in some cases). Once the task is accomplished, the job is done. However, that is not the real case in L2 language pedagogy. Willis (1996) pointed out that

> task-based learning is not just about getting learners to do one task and then another. If that were the case, learners would probably become quite expert at doing tasks and resourceful with their language, but they would almost certainly gain fluency at the expense of accuracy. ....to promote constant

> learning and improvement, we should see it (to do a task) as just one component in a larger framework (p.40).

Willis (1996) outlined the framework of task-based learning as consisting of three phases: a pre-task phase, a task cycle and a post-task language-focus stage, and argued that in addition to exposure and use of language, it is necessary to provide certain kinds of form focused instruction, either teacher-led or student-centered after the task cycle. Similarly, task researchers (Ellis, 2003; Skehan, 1996a; 1998) found it necessary to include a post task stage and distinguish three major stages for the implementation of tasks (pre, during and post task stages). In particular, Skehan (1996a) suggested that post-task activities can not be neglected with the assumption that learners' foreknowledge of the post task activity would influence how they allocate their attention during an actual task. To elaborate the view that post-task activities can change the way in which learners direct their attention during the task, Skehan (1996a) proposed two phases of post-task activities, phase one in a micro sense covering activities such as public performance, analysis and testing etc, and phase two in a macro sense including activities such as examining task sequences, task progression and task grouping. In the two phases, phase one is evidently concerned with the increasing of the accuracy and complexity effects which appears an appropriate phase for focus on form.

Skehan & Foster (1997) examined the effect of the foreknowledge of a certain post-task activity (to redo the task publicly) on three tasks (a personal, a

narrative and a decision making task). As they predicted, foreknowledge of public performance did not influence fluency and complexity, but had a significant effect on accuracy in the decision-making task. Although only one task (out of three) showed a significant accuracy effect, the research results were still encouraging in that the study showed that first, a post-task activity can have its function when it is operationalised in a suitable task context; secondly, the clearest effect of a post task activity is on accuracy (an attractive index for the goal of focus on form). This supports the claim that a post task stage may be a desirable phase for promoting the effect of focus on form. In an attempt to account for the weak effect of the public performance condition in Skehan & Foster (1997), Foster and Skehan (forthcoming) proposed that the weakness of the effect may be due to the nature of the activity—that is, public performance seems to have a remote influence on all the participants since only a small number (not all) of the students would be selected to perform in public; with regard to individual differences, public performance may not be viewed as a threat to certain participants (especially the brave learners); and in such an activity, for learners with different orientations in their performance (complexity-oriented, fluency-oriented or accuracy-oriented), their attention may not necessarily be directed to form to achieve a higher accuracy. In a word, 'public performance' cannot represent the whole picture of post-task activity.

In their subsequent study, Foster & Skehan (forthcoming) adopted another

alternative post-task activity— "transcribing" to eliminate the disadvantageous influence of public performance: first, transcribing is an individual activity for *every* participant to reexamine their own task performance, as such there would not occur any remote feeling among participants; second, to be pushed to transcribe their own performance, learners are forced to pay attention to the form of their own language during the task, thus accuracy becomes a necessity for all types of learners (Foster & Skehan, forthcoming). The researchers divided the participants into experimental and control groups. The experimental group transcribed extracts of their own task performance as a post-task activity while the control group did no post task activity. Both groups performed two types of tasks: narrative and decision-making tasks in two counterbalanced task cycles. The results showed that (1) foreknowledge of transcribing as a post-task activity has a significant accuracy effect on task performance in both narrative and decision making tasks; (2) significantly greater complexity was found for the experimental group in the decision making task. (3) With regard to fluency, one of the measures, length of run, was significantly greater for the post-task condition in the decision making task. As compared with Skehan & Foster (1997), the results in Foster and Skehan (forthcoming) are clearly more positive and supportive not only for the accuracy effect, but also for complexity and part of fluency. It is hoped that "transcribing" is likely to lead to an overall language improvement in terms of fluency, complexity and accuracy. As far as task type is concerned, more

interesting findings were revealed suggesting that the decision making task is a more promising arena to demonstrate the experimental effects not only in terms of accuracy, but also of complexity. In contrast, narrative performance is more difficult to influence. These results may attract researchers to pay more attention to the essential features of task type and its interactional function with other task conditions (Skehan & Foster, 2001; Skehan, 2007).

Skehan & Foster's post-task research is pioneering work in the task-based research field and shows some encouraging results for further studies. In addition, the post-task phase has also attracted attention from second language pedagogy. From a teacher's perspective, Lynch (2001, 2007) investigated the variables of transcribing as a post-task activity in an L2 classroom setting. Based on observation of classroom learning, two limitations had been identified by Lynch (2001): first, learners were given too much emphasis on learner activity, but too little time for reflection on the language. They kept on doing activities with few opportunities to reflect on their language gains and defects. Second, given that only high-proficiency learners were aware of their language changes in activity (Lynch & Maclean, 2001), devising ways to help learners analyse their performance is necessary. In such a context, Lynch employed transcribing as a reflective noticing activity for classroom learning. In contrast to employing transcribing as an individual task in Foster & Skehan's (forthcoming) study, Lynch (2001, 2007) designed the activity to be conducted in pairs as *collaborative*

transcribing. In Lynch's (2001) study, learners were asked to transcribe collaboratively, then discussed and revised the transcripts, and submitted the revised transcripts to the teacher for further corrections and reformulation. The analyses of the process and product of these cycles suggested that collaborative transcribing and revising can encourage learners to focus on form in their output in a relatively natural way. Furthermore, the teacher plays an important role in this post-task intervention, especially in the improvement of vocabulary (Lynch, 2001).

In a more recent study, Lynch (2007) compared the effects of two different transcribing groups—student-initiated transcribing (students in pairs transcribed their own performance and then revised the transcripts) and teacher initiated transcribing (the teacher transcribed problematic extracts of learners' performance, and the transcripts with errors were given back to the pairs for their revision). The analyses of the subsequent performance showed that both procedures are manageable under normal classroom conditions, and suggested that the student-initiated transcribing was more effective in helping the learners to maintain higher accuracy in the highlighted forms which were revised by students themselves.

The previous studies, either from a researcher's or from a teacher's perspective, show encouraging results with regard to the effect of post-task activities. Comparing the findings of two post-task activities (redo the task

publicly or transcribing task performance) in Skehan & Foster's two studies (1997, forthcoming), we find that transcribing may be adopted as a more feasible, manageable and effective post-task activity in communicative language classrooms. Similarly, Willis & Willis (2007) also suggest that transcribing as an effective post-task pedagogical choice is appealing to both teachers and learners. Some teachers have already begun to employ this activity in their classrooms (Clennell, 1999; Mennim, 2003).

As compared with the great amount of research concerning pre- and during task conditions, research concerning the post-task condition is surprisingly scarce. However, as far as the focus on form effect is concerned, the post-task phase is an important and promising stage to be explored (Skehan, 2007). Given this small body of existing research, there is clearly a need for further research on the effects of post-task activity on learners' task performance. Based on the previous studies' findings (Foster & Skehan, forthcoming; Lynch, 2001, 2007; Skehan & Foster, 1997), the present research will adopt transcribing (conducted in various ways) as a post-task activity and explore its effects on an extended scale.

2.3 Motivation for the present study

The goals of the current study are (i) to investigate whether employing transcribing as a post-task activity has an effect on focus on form in task performance; (ii) to compare the effects of different transcribing conditions in

terms of individual work vs. pair work, revision vs. no revision, student involvement vs. teacher intervention; (iii) to explore whether there is a practice effect in learners' performance if several task sessions are employed, in other words, whether it might be possible to train learners to distribute attention selectively over longer period; (iv) to compare the differences between two task types—the narrative and the decision-making interactive tasks.

The above goals are derived from the following reasons:

a. For a better focus on form effect, in other words to orient learners' attention to language form, transcribing may be viewed as one of the best activities. Asking learners to transcribe their own or one another's speech is not a new idea. We can draw on a long tradition of using dictation and its contemporary variant, dictogloss, both of which feature complete or partial transcribing (Lynch, 2001). Research shows that learners can benefit from transcribing either in a collaborative dialogue or from the other's presentation (Clennell, 1999; Swain & Lapkin, 1995). If transcribing speech pushes learners to focus on form, it is assumed that it would also be helpful for learners to transcribe their own speech, and to focus on their own language. Willis & Willis (2007) claim that when transcribing or listening to their performance for a second or third time, "participants process the language slightly differently each time, thereby assuring more chances to notice different linguistic features and maybe to incorporate some of them in their subsequent writing or speaking" (p. 168). As a result, Willis &

Willis (2007) recommend post-task transcribing and analysing as a helpful type of language work for L2 classroom pedagogy and it has already been adopted by many L2 teachers in various countries.

Furthermore, post-task transcribing may be viewed as a form of task repetition. This sort of task repetition can allow learners to draw on their familiarity with content and task demands, and with the process of formulating the desired meanings (c.f. Levelt's Speaking Model), so that they are able to devote more attention to getting the language right (Bygate, 1996, 2001; Gass et al, 1999). In addition, post-task transcribing is different from other sorts of task repetition in that it makes learners' speech visible instead of merely audible, and the visible transcripts increase the chance of learners' noticing, remembering and producing revised forms (Clennell, 1999; Johnson, 1996, Lynch, 2001, 2007). As a result, to explore the effects of various transcribing activities would be a promising aspect in both SLA research and pedagogy, particularly in terms of channeling learner's attention to a focus on form.

The issue concerning the operationalisation of a post-task activity is also worth exploring here. In contrast to Foster & Skehan (forthcoming) in which transcribing took place after class and out of the classroom, Lynch's studies embedded transcribing into the classroom immediately after the task performance. Transcribing on the spot would be more manageable and under teachers' or researchers' control in case that some unexpected intervening variables appear in

the research condition. In the present study, transcribing takes place immediately after the task performance and still in the classroom.

b. For a deep understanding on the effect of post-task transcribing, various conditions have been taken into account. In the present study, transcribing is employed either individually or in pairs with the aim to explore whether there is a difference between these two conditions. In Foster and Skehan (forthcoming), transcribing was operationalised on an individual basis. That is, each participant was asked to transcribe the task performance (including interactive task performance) by him/herself. In contrast, Lynch (2001, 2007) adopted a pair work style which he called *collaborative transcribing*.

Which would be more effective, individual work or pair work? From a theoretical perspective, the Interaction Hypothesis (Long, 1983, 1996) and Socio-cultural Theory (Lantolf, 2006) are in support of pair work. Long emphasizes that in the interaction between learners (or learners and native speakers), negotiation of meaning, or negative feedback such as recasting, can facilitate second language learning. Socio-cultural Theory proposes that in the interaction between an expert and a novice, the expert would provide scaffolding for the novice to attune his language and to reach a higher level of target language. As such, both theories agree that interaction in pairs would facilitate second language development. However, some research shows that pair work does not necessarily mean collaboration, and that it is only when learners work

collaboratively that they create opportunities for language learning. Furthermore, there has been relatively little empirical research comparing pair work and individual work. Foster & Skehan (1999) compared the effects of different pre-task planning conditions (teacher-led, solitary and group-based). The finding shows that group planning did not have advantages for learners' performance, and solitary planners produced more complex and more fluent language. Some researchers (Storch, 2005; Storch & Wigglesworth, 2007) compared short reports produced by students working in pairs with those of students working individually and found that pairs produced more grammatically accurate texts. To date, no consistent results have been reported and it is interesting to explore the difference between individual work and pair work in post-task transcribing activities.

c. In the current study, the experimental students are further divided into revision and no revision groups. Swain (1995, 1998, 2005) claimed that modified output is likely to promote language learning. To push learners to revise or modify their own output may cause not only attention to form, but also an underlying cognitive comparison (Doughty, 2001) and then a reprocessing process (c.f. Swain's Output Hypothesis). In Lynch (2001, 2007), students were involved in revising and reformulating. They showed performance improvement as a result of this process. Given the small sample size in Lynch's study (n=16), no generalization could be made concerning the effect of revision in transcribing. However, it is assumed that revision provides the opportunity for explicit

feedback, both positive and negative, and that this is a central requirement for formal language learning (Schachter, 1988; Ferris, 2004). In teaching practice, some teachers reflect that "during the transcription phase, many trainees wanted to write what they wish they had said rather than what they had actually said" (Willis & Willis, 2007, p 173). In other words, without communicative pressure, learners tend to recall their memory for better forms, or to test their language hypothesis for new forms. Learners' attention may be naturally channeled to language forms. The above theoretical background and the practice findings are the starting point for the design of the involvement of revision. Participants in the present research are engaged either in reflective self revision or interactive peer revision or supplementary teacher revision. It is hoped that the involvement of revision and reformulation in various forms would give rise to a change in learners' performance especially in terms of accuracy.

d. The teacher's role has also been incorporated into the present study. In Lynch (2001), it was showed that the teacher has a role to play in providing post-task feedback, particularly in the area of vocabulary. When the effects of self-correction and teacher correction on transcripts were compared in Lynch (2007), it was suggested that the self-transcribing was more effective in helping the learners to maintain higher accuracy. The above results show different roles for teachers within the transcribing activity.

In task-based teaching, the role of teachers has been enhanced to a more

demanding position (Ellis, 2003). Unlike the fixed role of the teacher in the traditional classroom, the role of the teacher in a task-based lesson varies at different stages. The teacher will act as the initial advisor, then the monitor during task, and the language guide and chairperson at the post-task stage (Willis, 1996). Some empirical studies have showed how teachers can have an effective role without compromising the naturalness of the task (Samuda, 2001; Thornbury, 1997; Toth, 2008). In contrast, some researchers have argued that "interactive practice during pair or group work presents several advantages over teacher-fronted activities" (Ortega, 2007:182). Interaction between learners affords better opportunities for the expression of a wider range of meanings and functions than is typically possible with teacher-fronted interactions (Hall & Walsh, 2002). To date, it is still not clear about the role of teachers in a task-based approach. The present research aims to explore the effects of the teacher's role in transcribing— as compared to the role of individual or pair students' work, whether teachers' intervention in revision would bring about any differences in learners' performance.

e. In Lynch's (2001, 2007) studies, only one task—'role play' was involved. This may cause some problems in the generalization of the results to other task types in classroom settings. Task researchers have contributed fruitful findings concerning the influence of task types and task features on learners' performance (for review, see Bygate, Skehan & Swain, 2001; Ellis, 2003; Skehan, 1998, 2007).

It would be more revealing to adopt different task types in a single research to see the effects of different task types. In this vein, the present study follows Foster & Skehan (forthcoming) in employing two types of tasks: a decision making task and a narrative task, the former of which was shown to be the most promising candidate demonstrating experimental effects and the latter to be the most difficult one to reveal significant differences among the experimental group. In this way, the validity of the current study is enhanced in that it takes into account the effects of different task types; and the results of the study may therefore be more generalizable for other classroom settings.

f. In contrast to most previous task-based studies, in the present study, participants are engaged in a longer period—four task sessions with one week interval between each cycle. The adoption of the multiple task sessions is derived from the related empirical findings in previous research and skill acquisition theory in cognitive psychology and applied linguistics. In Skehan & Foster (1997), all the participants were visited *three* times to complete three tasks at one-week intervals. The results in this regard suggested there was a practice effect, with accuracy scores in the third week higher (although not significantly) than those from the first week.

According to Dekeyser (2007), practice refers to "the specific activities in the second language, engaged in systematically, deliberately, with the goal of development knowledge of and skills in the second language" (p 1). In cognitive

psychology, the role of practice is rooted in skill acquisition theories, for example, Anderson's ACT* Model. In this model, learning is viewed as a certain kind of skill learning and it implies the development from initial controlled processing to a mediating associative stage, and to a final automatic stage. Practice plays an important role in turning the controlled declarative knowledge into the automatic procedural one (Anderson, 1993). It is generally agreed that "reaction time and error rate decline gradually as a function of practice with a given task" (Dekeyser, 2007: 4).

In second language acquisition, according to Ellis (1993), the purpose of L2 practice is "to develop fully proceduralized implicit knowledge" of the target language (p.109). He claimed that practice is important for teaching pronunciation and formulaic knowledge but not for the teaching of grammar rules (Ellis, 1993). Further, researchers who are advocates of input oriented approach take the view that input practice leads to language acquisition, but output practice merely serves to improve fluency (Ellis, 1993; VanPattern, 2002). In contrast, other researchers argued that sufficient communicative practice is an essential ingredient of successful language learning (Van den Branden, 2007), and output practice does not simply serve to increase fluency, but it leads to language acquisition as well (Swain, 1985; Swain & Lapkin, 1995). Given the controversy concerning the effects of practice on either language acquisition or communicative competence, there is a need to explore practice effects in the current study in which the focus is

on the language development or acquisition in a communicative setting.

One of the ways to enhance the practice effect is to provide multiple chances for learners to repeat the same or similar tasks, that is, to get learners to be involved in several task sessions. The idea of multiple repetition is "an important element of the praxis of practice all over the world" (DeKeyser, 2007: 11), and it is gradually accepted in empirical research on second language acquisition (Bygate, 1996, 2001; Gass et al 1999; Lynch & Maclean, 2000, 2001; Van den Branden, 1997, 2007). In the present study, four task sessions are designed for participants to complete two types of tasks (narrative task and decision-making task) twice to see whether there would be a growth in task performance over longer periods.

An additional benefit for the adoption of multiple task sessions is that it makes the measure of the treatment effect more revealing. Unlike the early studies on tasks or on focus on form which typically measured the effects only once, immediately following the completion of the experiments, later studies include a delayed post-test that makes it possible to see the durability of a certain treatment (Ellis, 2008). However, according to Ellis (2008), the ideal measure of treatment effects is 'process testing" (i.e. a whole series of tests), "as only this is capable of showing the accumulated effect of instruction on developmental transition over time" (p.843). In the present study, learners are involved in four task sessions, and the performance in every session is first analysed independently and then

compared with other sessions between groups and within groups. This measure is similar to what Ellis (2008) termed as "process testing" to explore learners' performance changes over time.

2.4 Research Questions

In view of the foregoing literature review, the present study investigates the effects of various types of transcribing as a post-task focus on form activity on learners' oral performance. Specifically, the aim of the study is multifaceted. In the first place, the general effect of transcribing, as compared to the non-transcribing condition, is the major focus of the study. In the second place, the relative effectiveness of pair work and individual work is paid much attention to as well. Further, the effect of revision is investigated. At the same time, the role of teacher's intervention in transcribing is taken into account. Moreover, the different effects of two task types are examined. Finally, the practice effect of the engagement of multiple task sessions has been explored. The research questions that guide the study are the following:

1). Do the post-task focus on form activities have effects on learners' oral performance?

2). Are there any different effects on the individual transcribing groups and on the pair transcribing groups?

3). Are there any different effects on the transcribing-only groups and on the

transcribing-and-revision groups?

4). Are there any different effects on the three revision groups? In other words, are there any differences between the teacher revising group and the student revising groups?

5). Are there any different effects of the two task types (i.e. narrative and decision-making tasks) on learners' oral performance?

6). Is there any practice effect as a result of the involvement of multiple task sessions?

2.5 Task Performance measures

Prior to the operation of the present study, it is necessary to note an important issue in task-based research: how task performance is measured. In the literature, task performance measures vary greatly among task researchers (Ellis, 2003; Ellis & Barkhuizen, 2005; Skehan, 1998). To a considerable extent, the different measures of task performance adopted in different studies have reflected researchers' theoretical positions (Skehan, 2003). Interactionists focus on the indices for the negotiation of meaning, such as clarification requests, confirmation checks, comprehension checks, recasts and uptakes (Long, 1983, 1996; Lyster, 1998, 2004; Mackey & Philip, 1998; Russell & Spada, 2006); Researchers based on the sociocultural theory prefer to the measures of interactive and language involvement, such as language-related episodes (LREs) and turns (Lowen, 2005;

Swain, 1995, 2005; Swain & Lapkin, 2001; Williams, 1998, 2000).Working in the framework of the cognitive approach, researchers distinguish oral performance in the following directions: fluency, accuracy, complexity, and recently the lexical aspects of performance (Foster & Skehan,1996; Robinson, 2001; Skehan, 1998; Skehan & Foster, 1997,1999, 2005; Yuan & Ellis, 2003).

Judged from its conceptual framework and rationale, the present study can be clearly located within the cognitive approach group. Thus, it follows the conventions adopted by a large number of cognitive task studies (Foster & Skehan, 1996; Skehan & Foster, 1997, 1999, 2005; Robinson, 2001; Yuan & Ellis, 2003) to focus on the measures of complexity, accuracy, fluency and lexical performance. However, even for the same construct of task performance (e.g. complexity), different measures are used across studies in the literature (Ortega, 2003). It is, therefore, needed to review the literature concerning the measure of task performance prior to the selection of performance measures and its application to the present study.

2.5.1 Conceptual background

In the traditional sense, some researchers argue that the three components of language performance (i.e. fluency, complexity and accuracy) reflect two different aspects of language processing: language representation (also known as declarative knowledge) and language access (also known as procedural

knowledge) (Wolfe-Quintero et al, 1998). In this view, complexity and accuracy correspond to the language representation of L2 learners. Particularly, complexity reveals "the scope of expanding or restructured knowledge to target language norms", accuracy shows "the conformity of L2 knowledge to target language norms" (Wolfe-Quintero et al, 1998, p 4). In comparison, fluency reflects how learners control the access of the language, "with control improving as the learner automatises the process of gaining access" (ibid, p.4).

From an information processing perspective, Skehan (1996a, 1998) argues that the three aspects—fluency, complexity and accuracy—all pertain to L2 language representation, but language representation consists of dual systems, and also different aspects of language performance draw on different linguistic systems: the exemplar-based and the rule-based systems. To achieve fluency, learners make use of the exemplar-based (or memory-based) system which consists of formulaic chunks for an increased production speed. In contrast, accuracy and complexity require learners to draw on the rule-based system in which the abstract rules can be used to create a variety of utterances/sentences. According to Skehan(1996b, 1998), in task-based language teaching and learning, fluency refers to real-time language production without undue pausing or hesitation which occurs when learners take the meaning as the primary concern to get the task done; complexity reflects learners' willingness to try the interlanguage structures that are 'cutting edge' and elaborated; and accuracy refers to "how well

the target language is produced in relation to the rule system of the target language (Skehan, 1996b, p23). Based on the different primacies of the three aspects, Skehan proposed that in language production, there would be an initial contrast between meaning (fluency) and form (complexity and accuracy). Further, the nature of complexity (i.e. restructuring of language) and of accuracy (i.e. control of language) can represent another contrast between them, although both complexity and accuracy are relevant for the rule-based system and viewed as the formal aspects of performance (Skehan, 1996b, 1998).

Independent of the above syntactic aspects, lexical performance has attracted attention in a relatively small number of task-based studies (Bygate, 1996; Foster, 1996; Foster & Skehan, 1996; Gass et al, 1999; Robinson, 1995, 2001; Skehan, 2009). Although vocabulary is now considered integral to every aspect of language knowledge (Daller, Milton, & Treffers-Daller, 2007), lexical performance has not received sufficient attention in the task-based research until recently. In their research concerning the measures of L2 development, Wolfe-Quintero et al (1998) include the lexical aspects as independent of fluency, accuracy and grammatical complexity and argue that lexical richness constitutes another major area for language complexity. According to Wolfe-Quintero et al, lexical complexity is manifested in terms of the range (lexical variation) and size (lexical sophistication) of L2 vocabulary. In a more recent book, Daller et al (2007) proposed the multiple measures of different aspects of lexical performance and

viewed lexical richness as an umbrella term covering lexical diversity, which is "the variety of active vocabulary deployed by a speakers or writer" (Malvern & Richards, 2002: 87), lexical sophistication (the number of low frequency words) and lexical density (the ratio of content and function words) (Daller et al, 2007:13).

In the present study, both the various syntactic aspects and lexical performance are taken into account. The following section reviews different measures of the above aspects in the literature and describes how those aspects of task performance are measured in the current research.

2.5.2 Measures of fluency, complexity, accuracy and lexical performance

Speech fluency is of such a multifaceted nature that it manifests and reflects not only underlying speech-planning and thinking processes but also the process of speech production, the phenomenon of hesitation and the temporal dimensions of speech (Freed, 2000). Lennon (1990) viewed oral fluency measures as two types: measures of temporal aspects, such as words per minute or pause length, and measures of dysfluencies, such as repairs. In the literature of SLA research, a comprehensive cluster of fluency-related measures were used including amount of speech, rate of speech, unfilled pauses, frequency of filled pauses, length of fluent speech runs, repairs (including repetitions, reformulations/false starts, corrections and partial repeats), and clusters of

dysfluencies (Koponen & Riggenbach, 2000; Freed, 2000). However, research shows that not all the fluency measures are necessarily effective in distinguishing the different oral proficiency levels, that is, not all the measures are perfectly reliable in oral data analysis (Freed, 2000). Pauses, for instance, may reflect either time required to focus on a new thought or time required to put a thought into words (Lennon, 1990). Thus, "the presence of pauses is not exclusively associated with a lack of fluency in a second language" (Freed, 2000:256). In the present study which concerns the effects of post-task focus on form activities, fluency, with a primary focus on meaning (Foster & Skehan, 1996), is not the major focus of performance measures. Thus, only some of potential fluency measures are adopted in the present study as follows. Given the multi-functionality of pauses in L2 speech, pauses are not included in the fluency measure. The adopted fluency measures are:

a. Speech rate: refers to how fast and dense the produced language is in terms of the time units. It is calculated on the basis of the number of (nonrepeated) words per minute;

b. Repair fluency: including the following categories: reformulation which suggests a decision to rephrase the form syntactically or morphologically, replacement which reflects the change of vocabulary in speech, false starts which occur where an utterance is begun and then is abandoned, and repetition which means the repetition of a word or a string of words (Foster, Tonkyn &

Wigglesworth, 2000; Skehan, coding manual).

c. Filled pauses: refers to the non-lexical fillers in the speech.

The above three kinds of fluency measures reflect both the temporal measures and the dysfluency aspects of oral performance which may reveal an overall picture of participants' fluency.

The following performance aspects are all form-related which are the main concerns of the present focus-on-form study. The measure of complexity is closely related with the various ways of segmenting language performance into units. Segmenting the language performance based on T-units is widely used in the analysis of written language. Hunt (1965) first developed the T-unit to measure children's syntactic maturity in writing, and defined the T-unit as a "minimal terminable unit" consisting of a main (i.e. independent clause) plus any subordinate clauses. Later, the definition of T-unit was adopted in the studies of spoken language as well. However, the T-unit has been criticized for the following reasons: first, the definition of T-unit only includes subordination, but not coordination, thus inappropriately dividing coordinations into different T-units (Bardovi-Harlig, 1992); second, the T-unit, originally developed for the analysis of written language, is not suitable for the analysis of spoken data which consist often of many fragments and elliptical sentences (Tarone, 1985; Foster et al, 2000). Alternatively, for spoken language analysis, some researchers adopted the communication unit (C-unit). The C-unit refers to not only grammatical

independent predications but answers to questions which lack only the repetition of the question elements (Loban, 1966 cited in Foster et al, 2000). For example, according to Loban, "Yes" can be viewed as a whole unit of communication when it is an answer to a question such as "have you ever been sick?" However, Foster et al (2000) pointed out that the definition of C-unit seems to exclude "elliptical constructions which arise within a speaker's turn rather than link to an interlocutor's question" (p.361) such as the topical noun phrase. It seems, therefore, that neither T-unit nor C-unit is suitable for the analysis of spoken data.

Foster et al (2000) proposed the "analysis of speech unit" (AS-unit) which can be applied reliably to spoken language. "An AS-unit is a single speaker's utterance consisting of an independent clause or sub-clausal unit, together with any subordinate clause(s) associated with either" (Foster et al, 2000: 365). An independent clause will be minimally a clause including a finite verb. As an analysis unit for spoken data, AS-unit also includes independent sub-clausal units which consist of either one or more phrases which can be elaborated to a full clause or a minor utterance which will be defined as irregular sentences or nonsentences, such as "thank you" "oh poor woman". (Foster et al, 2000:365-366). In this way, in contrast to the T-unit, fragments which are common in speech are included in AS-units. In addition, the AS-unit takes other features of spoken data into consideration, such as dysfluency features, topicalization, interruption and scaffolding.

The present study adopts the AS-unit as the segmentation unit for the measure of complexity. There are various types of grammatical complexity ratios. One type is the general complexity measure (clauses per production unit) which considers the proportion of all clause types to a larger unit. Another type is the dependent clause measure (dependent clauses per clause or per production unit etc) which considers the relationship between dependent and independent clauses. A third type is the coordination measure (coordinate clauses per clause or production unit etc) which considers the relationship between the coordination and independent clauses (Wolfe-Quintero et al, 1998). As compared to the dependent and coordination measures, a general complexity measure has been used in a larger number of studies, albeit with mixed findings (Wolfe-Quintero et al, 1998). The present research follows those previous studies to use a general complexity measure. In particular, complexity is measured by dividing the total number of clauses by the total number of AS-units. One crucial issue in the measure of grammatical complexity in spoken data is how to deal with fragments by researchers. As Foster et al (2000) proposed, fragments are included in AS-units. Similarly, following Bardovi-Harlig & Bofiman (1989), in the present study, a clause also includes sentence fragments with no overt verb. The more average number of clauses per AS-unit, the higher the complexity score (Skehan & Foster, 2005, Tavakoli & Skehan, 2005).

In addition to the AS-unit complexity ratio, the mean length of AS-unit is

considered as well. Although some researchers viewed the mean length of production unit as a fluency measure (Wolfe-Quintero et al, 1998), it is accepted among most researchers that the mean length of production unit is a measure of complexity (Iwashita, 2006; Norris & Ortega, 2008; Ortega, 2003) and the measure has been sufficiently investigated across studies (for a detailed review see Ortega, 2003).

As Norris & Ortega (2008) proposed that since syntactic complexity is multi-faceted, measures of complexity need to be multivariate. In past studies, the length of production units (e.g. T-units) and the number of clauses per T-unit are found to be the best ways to predict learner proficiency and also had a significant linear relation with independent oral proficiency measures (Iwashita, 2006; Ortega, 2003). To adopt more varied and valid measures of complexity for spoken data, the present research, therefore, uses the two effective measures of syntactic complexity—the mean length of AS-unit and the number of clauses per AS-unit.

Studies of second language acquisition have used various measures of accuracy. Some researchers measure accuracy by looking at the target-like usage of specific language features, such as past tense morphemes, plural morphemes, the number of correct pronouns or correct definite articles (Crookes, 1989; Cumming & Mellow, 1996; Ellis, 1987; Ortega, 1999; Wigglesworth, 1997). However, it has also been argued that such specific measures are less sensitive to detecting differences between experimental conditions (Bygate, 2001; Foster &

Skehan, 1996; Skehan & Foster, 1999).

The majority of researchers measure language accuracy by taking general errors into account. Two approaches are used in the literature. One is to focus on whether a structural unit (e.g. clauses, sentences, T-units or AS-unit) is error-free. Typical measures are the number of error-free T-units per T-unit, or the number of error-free clauses per total clauses, or the number of error-free clauses per T-unit (Wolfe-Quintero et al, 1998). Error-free measures have also been criticized for the following problems: first, they are concerned with the quantities of error-free strings, but not the quality (Polio,1997), in other words, this error-free measure does not reveal what types of errors are involved (Wolfe-Quintero et al, 1998); second, this measure does not reveal how errors are distributed within a unit, because a unit containing a single error is treated equally as a unit containing multiple errors (Bardovi-Harlig & Bofman, 1989); third, this error-free measure does not take the length of the analysis unit into account. Then, a high error-free ratio would be misleading when learners produce a large number of short but accurate analysis units (Skehan & Foster, 2005).

In view of the above criticism, researchers have developed other ways concerning accuracy measures. Some have proposed the calculation of errors in relation to production units (such as the number of errors per word, or per T-unit etc) (Bardovi-Harlig & Bofman, 1989; Fischer, 1984; Kepner, 1991). This approach is concerned with the quantities of errors, and it is better in

distinguishing production units with one error or more than one error than the error-free measure does. Further, Bardovi-Harlig & Bofman (1989) propose the use of clauses rather than T-units in the measure of accuracy in order to eliminate complexity as a factor. Other researchers improved the error-free measure to be more sensitive to the length of error-free clauses (Skehan & Foster, 2005). In addition to the ratio of error-free clauses, Skehan & Foster suggest calculating the error-free proportion in different clause lengths, that is to calculate the error-free proportion among three-word clauses, then among four-word clauses, and so on. Ideally, this would reveal a cut-off point, beyond which the participant cannot produce the correct clauses to meet a required criterion level. As for the criterion level, they suggest the 50%, 60% and 70% criterion levels be the likely candidates used in the task-based research (Skehan & Foster, 2005).

Bearing in mind the defects of the different accuracy measures and the corresponding developments, the current study takes the following accuracy measures collectively: first of all, the error-free clause ratio is calculated by dividing the number of error-free clauses by the total number of clauses without the interference of AS-unit segmentation; in addition, the error-free clause ratio is further calculated in different clause lengths, and a 70% criterion is adopted in the study to discriminate various accuracy levels of participants; further, to eliminate the defects of error-free measures, errors per 100 words is counted. Errors are defined as any deviation from the standard in terms of morphological, syntactic

and lexical aspects.

With regard to lexical performance, many commonly used measures have been based on the ratio of different words (types) to the total number of words (tokens), known as the Type-Token Ratio (TTR). The TTR has been criticized because it is sensitive to text length (Wolfe-Quintero et al, 1998). There is a negative relationship between a type/token ratio and sample size, that is samples with larger numbers of tokens give lower values for TTR and vice versa, because the longer samples of language that are produced, the more of the active vocabulary is likely to be included and the available pool of new word types that can be introduced steadily diminishes (Malvern & Richards, 2002). Given that most task performances of second language learners are fairly short, the TTR for L2 learners lexical performance poses acute problems (Skehan, 2003). According to Read (2000), the TTR should be regarded just as a single measure of lexical diversity. Nation (2007) strongly argues that since vocabulary knowledge is multi-dimensional, it is necessary to adopt a set of complementary measures that tap into different aspects of vocabulary knowledge and give a more complete picture than a single measure can. In attempt to arrive at such multiple measures, researchers proposed the concept of "lexical richness" which includes different aspects of vocabulary use, such as lexical variation (i.e. lexical diversity), lexical density and lexical sophistication (Daller et al, 2007). In their synthesis research, Wolfe-Quintero et al (1998) find that lexical variation and lexical sophistication

are related to language development, but not lexical density. In the field of vocabulary assessment and measures, researchers pay more attention to lexical sophistication and lexical variation than to lexical density (see Daller et al, 2007). Following the past studies, the present research takes lexical variation and lexical sophistication as the methods of measuring lexical performance.

Lexical variation (also lexical diversity) refers to the variety of learners' vocabulary. As an alternative to the type/token ratio, vocd[1] was developed by Malvern & Richards (1997, 2002) which assumes another functional relationship between the number of types and tokens, thus providing a new measure of vocabulary diversity referred as the D value. Put simply, D provides an index of the extent to which the speaker avoids the recycling of the same set of words. A lower D suggests a larger tendency to return to a set of words by the speaker. Thus, D is a text-internal measure of lexical performance (Skehan, 2009). D has been shown to be superior to previous measures in both avoiding the inherent flaw in raw TTR with varying sample sizes and in comparing different speakers on the basis of the number of types and tokens they produce irrespective of the length of texts (Malvern & Richards, 2000). However, D is not perfect in that as a type/token-based measure it does not take into account the frequency of a word. A common word and a rare one have the same weights in the measure of lexical variation.

[1] Vocd is the command in CLAN for the computation of D.

Unlike D and other type-token based measures, the measure for lexical sophistication is based on the notion of frequency. Lexical sophistication is measured by determining the number of low frequency words in a text (Read, 2000). P-Lex, for example, as a computer program can do this automatically by dividing a text into ten word chunks and then calculating the number of infrequent words in each ten word chunk (Meara & Bell, 2001; Skehan, 2009). P-Lex is based on a probability distribution (Poisson distribution) that is taken as a model for the occurrence of rare or difficult words, and a Poisson distribution is reflected by a single parameter Lambda value. To define the rare or difficult words, a word frequency list is needed in P-Lex. As Daller & Xue (2007) pointed out, the word lists for P-Lex "have to be chosen carefully and have to be adapted to the specific task" (p.164). For the narrative and the decision-making tasks used in the present study, P-Lex was rewritten by Skehan with the reference to the British National Corpus spoken component. In this modified program, the word list is lemmatised. Files of task-specific words are compiled to enable words to be temporarily defined as easy. In addition, a cut-off value of fewer uses than 150 per million words is used to define words difficulty (Skehan, 2009). "This value seemed to be most effective in producing a good range of discrimination" (Skehan, 2009: p.110).

The two kinds of measures for lexical variation and lexical sophistication are complementary to each other—the frequency-based measure of lexical

sophistication is a text-external measure and reveals "the access to and deployment of more rare or more difficult, more precise vocabulary" (Richards & Malvern, 2007:84), while the type-token based measure of lexical variation reflects the "access to a wide range of vocabulary, and by inference, its skillful use" (Richards & Malvern, 2007:84). The present study adopts the two measures of lexical richness in hope to give a complete picture of lexical performance in different experimental groups.

In the current research, task performance is measured in terms of fluency, complexity, accuracy, and lexical performance (i.e. lexical variation and lexical sophistication). To tap into the different features of a performance construct, multivariate measures are adopted for the same construct. Those measures may be independent of each other, but they are complementary to each other to provide an overall picture of learners' task performance.

Chapter Three Research Methodology

3.0 Introduction

This chapter starts with the description of pilot studies undertaken prior to the main study. The pilot studies determined a) the type of tasks and topics that were to be used in the main study and b) the operational procedures of the main study. Then, the chapter presents the detail of the main study, including the participants in the main study, the experimental tasks adopted, the research design, the setting of the experiments, the experiment procedures followed, the procedures for data transcribing and coding, and finally the methods for data analysis.

3.1 Pilot Studies

Two pilot studies were carried out: the first one was designed for the selection of tasks and topics which would be used in the main study, and the second one is to pilot the feasibility of procedures in the main study, at the same time, to identify the intervening factors in the procedures which should be controlled or removed from the main study.

3.1.1 Pilot Study I

The present study follows the previous literature (Skehan & Foster, 1997; Foster & Skehan, forthcoming) in adopting two different task types: narrative

tasks and interactive tasks. The narrative tasks are video retelling tasks, in which a number of different episodes of *Tom and Jerry* are used in different narrative task sessions. The interactive tasks are discussion tasks, in which participants discussed several problematic issues in various sessions. It is, therefore, necessary to ensure that the video episodes and the discussion topics are of equivalent difficulty so as to remove the intervening effects of different task materials and task topics in the task performances. Thus, one of the preliminary concerns of the pilot study is to choose the task materials (for narrative tasks) and task topics (for interactive tasks) which are supposed to be equal or similar in terms of the degree of task difficulty. The first pilot study is designed for the selection of task materials and topics.

3.1.1.1 Selecting task materials for narrative tasks

Following the previous studies which also adopted cartoon episodes from *Tom and Jerry* (Bygate, 1996, 2001; Gass et al, 1999), the present study adopted several criteria for the initial selection of the episodes:

a. there are no conversations in the stories, thus no extra pressure on learners' listening comprehension will occur to distract learners' attention on tasks;

b. the selected stories do not have any specific culture background which may cause certain cultural bias or misinterpretation among the audience;

c. there are fewer than four characters in each story and the plots are

similarly simple so that no comprehension barrier may hinder learners' task performance.

d. the selected episodes are of similar length so that they require more or less the same attention from the audience and the study can be conducted under the similar conditions.

Based on the above criteria, four episodes out of 20 were initially selected by the researcher. Then, the four episodes were rated by five tertiary English teachers with regard to the level of difficulty concerning the background knowledge required, the difficulty in terms of comprehension of the episodes and in terms of retelling the episodes, the amount of vocabulary to retell the stories and the amount of syntax in retelling etc. The rating scheme is attached in Appendix I. As a result, three episodes which got similar scores in all aspects were selected for further piloting. The reason to choose three episodes is that there are a training session and two narrative task sessions in the major study which need three episodes to be available.

In the first section of pilot study I, two participants were invited to do the narrative tasks. The two participants were second year university students (one male and one female). The participants in the pilot and in the major studies were from the same participants' pool (i.e. students of the same year and same major in one university). They performed the narrative tasks separately but in different episode orders to avoid the influence of task sequence. Their narrative task

performances were recorded for further analyses by the researcher.

After the narrative task performances, the participants were interviewed for their feedback generally concerning the difficulty of the video comprehension and the difficulty of the oral performance. They were free to choose either their first language (Chinese) or English to answer the questions. Both of them chose to answer the questions in English. As far as the narrative task was concerned, the interview covered the following questions:

1. When you were watching the three episodes, which one do you think was the most difficult one to understand, and which one was the easiest one? Or do you think they were almost equal for you to understand?

2. When you were retelling the stories, which story do you think was the most difficult to organize the language? Or were they almost similar for you to retell the stories?

3. Have you met any problems in understanding the stories? If yes, which plot or which scene is challenging for you to understand?

4. Have you met any problems in retelling the stories? If yes, which kinds of problems are they? Are they problems of vocabulary, syntax or meaning?

5. Have you ever experienced this kind of English practice in your secondary school or university? If yes, could you talk about your English class?

The participants' responses on the above questions can be summarized as the following: first, both participants regarded their first narrative task as the most

difficult one in terms of both comprehension and retelling, and the other two are similarly easier than the first narrative task. Thus, they felt more nervous in the first task, but more relaxed in the following tasks. However, with the counterbalanced arrangement of the episodes, the two participants watched and retold different episodes as their first narrative task. This suggests that the sequence of task materials is of importance to influence the narrative task performance. Secondly, they did not find much difficulty in comprehending the stories. However, they treated the retelling task as a challenging one since they were worried about the language they used in the performance. They encountered problems with regard to both syntax and the lexicon. Thirdly, they had experienced a similar type of retelling task—picture describing tasks. But they preferred the video retelling which they regarded as more interesting and easier for them to comprehend and organize the language in retelling. In addition, in their English class, they did not need to do further things after retelling, which may reveal a lack of post-task activities in their secondary or university English classes.

Upon the accomplishment of the first section of pilot study I, the narrative task performances were analysed in terms of accuracy (the ratio of error-free clauses), complexity (clauses of per AS-unit) and fluency (words per minute, i.e. speech rate). The descriptive results (see Table 3.1) show that for both participants, their first narrative task performance got the lowest scores in terms of all the

above aspects of task performance, while the other two performances show better results.

Table 3.1. Descriptive Results for Narrative Tasks in Pilot Study I

	accuracy			complexity			fluency		
	N1	N2	N3	N1	N2	N3	N1	N2	N3
participant A	0.32	0.48	0.51	1.21	1.42	1.37	43.12	50.15	52.17
participant B	0.41	0.52	0.53	1.17	1.23	1.30	41.21	55.34	51.48

In sum, in the first section of pilot study I, we can get the following suggestions for the major study:

a. The three selected episodes in *Tom and Jerry* are of similar degree of difficulty in terms of both comprehension and retelling. Thus, they can be adopted in the main study, one for the training session and the other two for the two narrative task sessions.

b. The sequence of task materials impacts greatly on video understanding and retelling. It is, therefore, necessary to use a counterbalanced design in the main study.

3.1.1.2 Selecting discussion topics for interactive tasks

In the second section of pilot study I, the same participants did two interactive tasks in a dyad. Initially, the interactive tasks were designed as discussion tasks. In particular, two discussion tasks were involved—discussion of

a scholarship award and a jury discussion on an accident.

For the scholarship award task, participants played as the members of a university scholarship committee. They discussed the applications of three outstanding students and then decided to whom the scholarship would be awarded.. For the jury discussion task, participants role-played as jury members and then decided which party or parties should be responsible for the accident. The information of the accident was adapted from a newspaper article. The topics for the discussion tasks are attached in Appendix II. Their interactive tasks were recorded for further analyses.

After the discussion tasks, the participants were interviewed with regard to their feedback on the tasks. The interview questions are the following:

1. When you were discussing the two topics, were you familiar with both of them?

2. Were they interesting or boring to introduce you to perform the discussion? Why?

3. Which one do you think is more difficult for you to carry out the discussion? Or are they similar for you to perform the discussions?

4. What aspect is the most challenging one for you to perform the discussion? (e.g. the content, the different perspectives, or the language)

5. What aspect is the easiest one for you to perform the discussion? (e.g. the content, the different perspectives, or the language)

6. Have you ever experienced such discussion tasks in your English classes? If yes, what did you do after discussion?

The participants' responses to the interview are summarized in the following: first of all, they treated the two topics as equally familiar but not equally interesting. In addition, the two topics were of similar difficulty for them to conduct discussions. However, they thought the most challenging aspect in the tasks was to develop the discussion further, as they cannot contribute different perspectives of ideas on the two topics, and they only provided one or two turns in the conversation then reached an agreement. As far as their previous experience is concerned, although they have met this type of discussion tasks in their English class, they always felt helpless during discussions, because they did not know how to develop the discussion. Furthermore, they did nothing after the discussion tasks in their English class.

The task performances reflected what the participants felt in the interview. In each discussion task, only one or two turns were contributed by each speaker. This simple conversation may not provide sufficient information for interactive task performance analyses in the present study. The selection of the two discussion topics was regarded as a failure. As a result, the interactive tasks need to be designed in another way in further pilot studies.

3.1.1.3 Selecting another type of interactive tasks

Given the topic failure in the second section of pilot study one, it was necessary to choose another type of interactive tasks in the hope that the newly-selected interactive task type could encourage participants to contribute as many turns as possible in task performance. Following previous post-task research (Foster & Skehan, forthcoming; Skehan & Foster, 1997), another type of interactive task—the problem letter task—was adopted to replace the above discussion tasks. While doing the problem letter task, the dyads acted as journal editors discussing various solutions for a tricky problem written in a letter. Different suggestions were expected to be provided and analyzed while finally the best one would be selected as the reply to the letter writer. Three topics to be included in the letters were selected from news journals which were all concerned with common problems among university students so as to ensure that participants were familiar with the problems and would contribute as many ideas as possible. The adopted problem letters are the following.

We have a daughter of 16 year old. When she was in junior middle school, she was excellent.

Last year, she began to be addicted to internet exploration. She made some net-friends through the internet. Recently she disappeared for two weeks. At last, we found her in a boy's house. She told us that she was going to marry that boy whom she got to know through the Internet. The boy is 12 years

older than her, and now is jobless and staying at home. We locked her in the bedroom and didn't allow her to go out. However, she escaped from the window to the boy's house. What shall we do now?

I have been in love with my girlfriend for one year. Now she is studying at a university and I work as a clerk in a company after my graduation from a technology institute. Being afraid of her parents' objection, she didn't tell her parents about our love until recent. Several days ago, she told them and quarreled with them. Her mother came to my company and asked me to break up with her, because I have a lower academic degree than her daughter. I am in despair. What shall I do now?

Our boy is 17 years old. He is excellent at both playing badminton and doing academic study. Now, our boy has been selected to be the member of national badminton youth team. We are now hesitating about whether to send him to the badminton team. To be a professional badminton player or to be an excellent scientist (he has a talent for physics)? It is hard to decide. What shall we do?

In the third section, the same participants performed the above problem-letter tasks, each in a dyad (see appendix III for the task instruction). Task performances were recorded for further analyses. Similar to the above two pilot sections, the

two participants received an interview after the task performances. Besides the questions presented in the section of pilot study I, the other interview questions concerned the comparison between the problem-letter tasks and the previous discussion tasks. In the interview, the participants reflected that firstly, they thought the three problem-letter tasks were of similar degree of difficulty for them to perform the discussion. We may assume, therefore, there would not be any effects of task topics on interactive task performance in the main study. Secondly, they preferred this task to the previous discussion tasks in that they could produce more ideas from different perspectives because the topics were more closely related with the university students themselves. Thirdly, they had encountered the same problems with regard to lexical and syntactic aspects of language as they did in the narrative and the previous discussion tasks. Fourthly, they had met certain types of decision-making tasks which were similar to this problem-letter task in their regular English classes.

Their task performances were analysed with regard to accuracy, complexity and fluency. The results show that there are no great differences among the three task performances. The results are presented in Table 3.2. Consequently, it was decided to choose the problem-letter task as a type of the decision-making tasks (i.e. interactive task) in the main study.

Table 3.2. Descriptive Results for Interactive Tasks in Pilot Study I

	accuracy			Complexity			Fluency		
	DM1	DM2	DM3	DM1	DM 2	DM 3	DM 1	DM 2	DM 3
participant A	0.46	0.50	0.54	1.51	1.53	1.60	63.14	65.42	61.53
participant B	0.53	0.52	0.57	1.48	1.72	1.51	73.25	69.81	65.36

3.1.1.4 Summary of pilot study I

The aim of the first pilot study was to select appropriate task materials and topics of similar or equivalent difficulty for different narrative and interactive task sessions. Three pilot sections were conducted and the results are summarized as the following:

a. As far as the narrative task is concerned, the sequence of video episodes has a considerable impact on task performance. Participants felt more nervous in the first narrative task than the following ones. As a result, their narrative task performance was worse in the first task than the other two. It was, therefore, decided to employ a training session and to provide a short period of planning time[2] prior to the main study on the one hand, and to adopt a counterbalanced design concerning the arrangement of the episodes and topics on the other hand.

[2] According to the planning literature (Mehnert, 1998), the provision of pre-task planning time between one and five minutes did not make differences in terms of accuracy and complexity which are the major concerns of the present post-task focus on form research. In pilot study II and the main study, to keep control of the intervening factor of pre-task planning, two-minute planning time is given to all the groups equally prior to the task performance.

b. The three selected episodes were of a similar degree of difficulty in terms of both understanding and retelling. As a result, they were adopted for the narrative tasks in the main study.

c. Given the failure of the scholarship award and the jury discussion tasks, the problem-letter tasks are more effective to introduce participants to elaborate on the selected topics. Furthermore, the three topics have similar impacts on task performance. They are employed as the decision-making task topics in the main study.

3.1.2 Pilot Study II

Upon the selection of the task materials and topics for the two task types, the second pilot study was carried out a) to pilot various types of post-task focus on form activities so as to choose the comparable ones for the main study, b) to see whether the major procedures were feasible so as to adjust the further design of the present research, and c) to explore any unexpected factors during the procedures so that they may be well controlled in the main study.

3.1.2.1 Procedures of pilot study II

In the second pilot study, three different post-task activities were designed: a) a teacher-led focus-on-form[3] activity—the teacher analysed the performance

[3] Prior to the pilot study, the preliminary focus of the research was not limited to the post-task transcribing only. On the other hand, various post-task FonF activities were considered so as to seek for those which can be comparable and operationally feasible in the main study.

recordings after class and then chose the major language problems to be explained prior to the next task session; b) post-task transcribing—after they had finished the tasks, the participants listened to the recordings and transcribed parts of their own task performance individually c) post-task transcribing plus revision—once they had accomplished the tasks, the participants transcribed parts of their own task performances individually and then revised the transcripts for a better version of the task performances.

One point should be noted with regard to the performance length to be transcribed at the post-task stage. In pilot study one, we found that the participants took six to seven minutes to retell a 6-minute video episode in *Tom and Jerry*, and the same participants needed more than 7 minutes to accomplish the interactive tasks. Since around three minutes are needed for an advanced English graduate to transcribe one-minute task performance, we may assume that the participants in the present study who are at intermediate level of English proficiency will need more time for the transcribing task (Skehan, 2007, personal communication). For reasonable time control, therefore, the two transcribing groups were asked to transcribe a three-minute extract from their task performance.

Twelve students from the same participants' pool of the main study participated the second pilot study. They were randomly arranged into three groups for the above different types of post-task activities respectively. During three consecutive days the groups performed a narrative task on the first day, an

interactive task on the second day and both another narrative and another interactive task on the third day. Based on the summary of pilot study I, a training session and a two-minute pre-task planning time period were provided for all the groups so as to make the participants familiar of the task procedures and to reduce the pressure on the subsequent task performance. The design for the Pilot Study II is presented in Table 3.3.

Table 3.3. Research Design for Pilot Study II

	Group I	Group II	Group III
Day One	a. training on the narrative task b. planning & narrative task 1	a. training on the narrative tasks b. planning & narrative task 1 c. post-task transcribing	a. training on the narrative task; b. planning & narrative task 1 c.post-task transcribing + revision
Day Two	a. teacher-led FonF b. training on the interactive task c. planning & interactive task 1	a. training on the interactive task b. planning & interactive 1 c. post-task transcribing	a. training on the interactive task b. planning & interactive task 1 c.post-task transcribing + revision
Day Three	a. teacher-led FonF b. narrative task 2 c. interactive task 2	a. narrative task 2 b. interactive task 2	a. narrative task 2 b. interactive task 2

(Note: the post-task activities are underlined for emphasis).

At the end of pilot study II, the participants had an interview to provide feedback about the procedures and the post-task activities. The following questions were included:

1. What do you think of the training session before the task? How did you feel before the task training on video retelling task? How did you feel after the training?

2. How did you treat the 2-minute planning time before the performance? What did you do during the planning time? On the last day, you were not given planning time. Were there any different effects on your performance with or without planning? Do you prefer planning or no planning before the task performance?
3. You have been involved in a certain type of activity after task performance, such as teacher's explanation of the language problems (for Group I), post-task transcribing (for Group II), and post-task transcribing plus revision (for Group III). Do you have any comments on the activity? Is it useful? For language improvement or for communicative ability? Would you please state the reason for your comments?
4. Are there any differences between the two videos or the two discussion topics?
5. You general comments about the three days' sessions.

The participants' responses are summarized as follows: first of all, participants reflected that the training time was helpful and important for the following task performances. During the training session, some basic strategies and skills for video retelling and problem solving tasks were provided by the teacher. According to the participants, they adopted most of the strategies in the subsequent task performance. In addition, the pre-task planning time was useful for them as well. During the planning time, mainly they prepared for the organization of the ideas in both tasks. All participants preferred to be given planning time.

As regards the three types of post-task activities, participants in different

groups commented on the post-task activities in which they were involved. For the teacher-led focus-on-form group, participants were familiar with the activity in which a teacher explained their errors or problems in written or oral tasks. However, the explanations were more general than specific when the whole class' errors were collected together and only some of them were analysed. In the pilot study, they thought the error correction and explanation were focusing on the common and typical errors which they had already known before. However, they found they made errors in the same way in the subsequent tasks.

For the post-task transcribing group, participants treated the transcribing activity mainly as a listening training activity. As a result of transcribing their own performance, the participants would like to pay more attention to their pronunciation in the following tasks, and at the same time, some attention was paid to aspects of language. However, they admitted that even though they focused on the language and made efforts to avoid the previous errors, they repeated most of the same errors in the following task performances.

For the post-task transcribing plus revision group, participants regarded this activity as useful in that when they revised their own transcripts, they focused most attention on the language itself. In contrast to Group II, none of the participants in this group treated the post-task activity as listening training. All were concerned with the language in the transcripts, and they managed to avoid certain errors in the following tasks. One point is worth mentioning here. The

participants from the three groups preferred the adoption of post-task activities (although in various forms for different groups) and viewed it as even more important than just asking them to do the tasks in the class.

As for the differences between task materials and task topics, participants reflected that the first narrative task was a bit more difficult than the second, whereas the two discussion topics were more or less of similar difficulty in terms of content elaboration and language requirements. This matched the results in the first section of pilot study one in which the participants viewed the first narrative task as more demanding than the second one. Given that no similar reflections were presented concerning the interactive tasks, it may be suggested that task sequence impacts greatly on the narrative tasks, but much less so on the interactive tasks.

In addition to the participants' feedback in pilot study II, the teacher's reflection afterwards was also taken into consideration. In the present research, the researcher acted as the teacher in both the pilot and the main study. Among the three different post-task activities, the teacher plays a more important role in the teacher-led focus-on-form group than in the other groups. After the completion of the second pilot study, the research suggested that it was rather challenging to cover all the errors which occurred in the task-performances for all the students involved. As a result, explanations about the most common or typical errors in the task performances were not as specific and effective as they might have been.

More important, the researcher (the teacher in the study) was not the regular teacher for the participants. In the literature some research (Ferris, 2003; Hyland, 1998; Montgomery & Baker, 2007) has identified various factors that may influence the effectiveness of teachers' feedback on errors. One such factor is concerned with a possible lack of sensitivity on the part of teachers to different contexts as well as to varying levels of need, ability, and other individual differences of students. That is the case for the teacher in the present study. Without good knowledge about the participants, the temporary teacher may not be qualified to conduct as effective a focus-on-form activity as the regular teacher might do. In other words, we may say that whether the focus on form activity is conducted by the regular or the temporary teacher appeared as an unexpected intervening factor in the study.[4] Based on the reflections from both the participants and the teacher (i.e. the researcher), it was decided, therefore, to adjust the design of the main study, that is, to replace the teacher-led focus on form group with several varieties of post-task transcribing groups, since at the same time, the second pilot study showed that the post-task transcribing which is learner-involved rather than teacher-led, can be operated smoothly without the unexpected intervening factors mentioned above.

[4] In the main study, around 100 participants were from different majors and classes. So it was not reasonable to invite the different English teachers to teach different groups in the main study. In effect, the involvement of multiple teachers in a single study would also create certain intervening factors, such as the different preferences of teachers.

3.1.2.2 Summary of pilot study II

For the second pilot study, the major concern was to pilot the operation of various post-task activities and to explore the feasibility of the procedures so as to remove any potential intervening factors in the main study. The following suggestions are relevant for the main study:

a. The major procedures of the study are feasible and under reasonable control. In addition, post-task transcribing seems to be effective and welcomed by the participants. In the main study, it is feasible to develop post-task transcribing into different conditions to explore the different effects of various post-task transcribing activities.

b. The adjustment of the design in the main study is related to the replacement of a certain type of post-task activity. In particular, the teacher-led focus on form activity at the post-task stage, which tends to be easily influenced by certain intervening factors as mentioned above, will be replaced by a variety of post-task transcribing activities in the main study.

c. The performance to be transcribed is set to be 3 minutes in length in both tasks. The advantages are two-fold: first, a 3-minute extract provides sufficient task performance to be transcribed which can give rich language exposure to the participants. Second, a reasonably equal period of time, neither too long nor too short, can be provided

to transcribe different performances. Thus, it can be ensured that the time distribution of each task session will be managed effectively in the main study.

In sum, the two pilot studies are beneficial in that they pave the way for the main research by teasing out the intervening factors, such as the effects of task materials and task topics, the negative effects of outside teachers etc. Furthermore, the findings of the pilot studies ensure the feasibility of the general procedure by taking all relevant details into consideration prior to the main study.

3.2 Participants

Initially, one hundred and twenty-two students were involved in this study. In final analysis, to avoid the influence of topic familiarity and to ensure data completeness, the following oral data were excluded if a) the participants had ever watched the cartoon episodes before; b) the participants had not attended all the research meetings. Finally, ninety-six participants were included in this study. All of them were second-year university students from a south China comprehensive university. They were non-English majors who were majoring in history, Chinese, physics, chemical engineering and computer science. Forty one were female and fifty five were male. They had been studying English for 7-10 years. The students participated voluntarily and were not paid. As a reward for their participation, an

analysis report concerning their performances was given when the study had been accomplished and the participants were interested in their personal results. The participants were divided into five experimental groups and one control group randomly only according to the timeslots which would be available for them to attend the study.

Prior to the main study, an important issue was the comparability of the six groups regarding the English proficiency. To explore this, a proficiency cloze test was administered among the participants. In language testing, cloze tests have frequently been adopted as a valid and reliable instrument to assess overall language proficiency (Brown & Rodgers, 2002). Three cloze passages (see Appendix IV) were adopted from nation-wide standardized China College English Test Band 4 (CET-4) Database. CET-4 is a nation-wide standardized test for all the non-English majors in China's universities. As a result, the adopted cloze test passages were supposed to be valid and reliable. The three passages had been applied in different previous CET-4 tests which were administered several years ago. So it was assumed that no participants had done the same tests before. In the answer paper, the participants were asked to indicate whether they had done any of the passages before. It showed that it was the first time for the participants to be tested on the three passages. Table 3.4 shows both the distribution of participants and the means of the proficiency test for each group. A one way ANOVA showed no significant difference among the six groups on the cloze test, $F(5, 94)=.628$,

p=.679. On this basis, we can conclude that the six groups were comparable with regard to English proficiency. This established the pre-condition for the operation of the study.

Table 3.4: Participants Distribution and Proficiency Means for Six Groups

	individual transcribing group	individual transcribing & revision group	pair transcribing group	pair transcribing & pair revision group	pair transcribing & teacher revising group	control group	total
No.	16	16	16	16	16	16	96
Means	38.94	37.88	38.82	39.11	39.19	40.94	39.14
SD	6.03	5.12	4.67	4.78	5.06	4.419	4.993

3.3 Experimental Tasks

Following previous related research (Foster & Skehan, forthcoming; Lynch, 2001, 2007; Skehan & Foster 1997), the present study used only a narrative task and a decision-making task, with the exclusion of the personal task. The reasons lie in the following. In Skehan and Foster's series of studies (1997,1999,2001), they found that a narrative task is a most challenging task for the second language learners to perform on the one hand, and for the researchers to demonstrate the experimental effects on the other hand. In contrast, a decision-making task was the only task that did generate a significant accuracy effect in previous studies, while the personal task appeared at the intermediate point to display that effect. In focus on form studies, the accuracy effect is a desirable indicator of the treatment

effect (Doughty & Williams, 1998; Norris & Ortega, 2000, Skehan, 1998). Thus, for the present study, as Skehan & Foster suggested, we can assume if a significant accuracy effect would be manifested with a most challenge task-- a narrative task, then we are likely to expect a similar effect with personal information tasks (Foster & Skehan, forthcoming). In order to see the relative long-term effect of post-task activities, and of task practice, the present study adopted two narrative tasks and two decision-making tasks for real task performance, and one narrative task and one decision-making task for task training in the orientation period.

For greater methodological control, the four treatment tasks were carefully arranged in order to tease out the intervening influence of task type and task order. The two sub-groups under each treatment group were assigned the same tasks in reverse orders to counterbalance that intervening effect of task sequence. Table 3.5 shows the task arrangement.

Table 3.5. Task Arrangement

	Sub-group 1	Sub-group 2
Task Cycle One	Narrative Task a	Decision-making Task b
Task Cycle Two	Decision-making Task a	Narrative Task b
Task Cycle Three	Decision-making Task b	Narrative Task a
Task Cycle Four	Narrative Task b	Decision-making Task a

In the narrative task, the participants described the stories of cartoon episodes after watching. In the decision-making task, the participants in dyad or in a group of three acted as the editors of a magazine problem column. They discussed the problem in a letter written to the magazine and agreed upon some best advices for the writer. Each letter described a certain tricky personal situation that did not have simple or obvious solutions (Skehan & Foster, 2001).

3.4 Research Design

A 5×2×2 research design was employed. The first independent variable, the post-task transcribing variety, is a between subject factor with five levels: 1) individual transcribing only, 2) individual transcribing and revising, 3) pair transcribing only, 4) pair transcribing and pair revising, 5) pair transcribing and teacher revising. The second independent variable, task type, is a within-subject factor having two levels: the narrative task and the decision-making task. The third independent variable, task session, is a within-subject factor with two levels: the first and the second sessions for either narrative or decision-making tasks. This design allows for the between subject comparison for the variety of the

transcribing conditions and within-subject comparisons for task practice as result of several task sessions and for the different task types. The dependent variable is the oral performance which was measured in terms of fluency, accuracy, complexity, and lexical performance. Table 3.6 shows the design of the study.

Table 3.6. Research Design

Group / Cycle	Individual Transcribing (Group 1)		Individual Transcribing & Revising (Group 2)		Pair Transcribing (Group 3)		Pair Transcribing & Revising (Group 4)		Pair Transcribing & Teacher revising (Group 5)		Control (Group 6)	
	Sub-group 1	Sub-group2	Sub-group 1	Sub-group2	Sub-group 1	Sub-group 2	Sub-group 1	Sub-group2	Sub-group 1	Sub-group 2	Sub-group 1	Sub-group 2
Orientation	1) To make participants be well informed of the experiment, mainly task training including how to perform the tasks, and how to transcribe and revise. 2) To administer a proficiency test to ensure that the groups are comparable with regard to mean English proficiency.											
cycle 1	Na[1] Post-task 1[2]	DMb Post-task 1	Na Post-task 2	DMb Post-task 2	Na Post-task 3	DMb Post-task 3	Na Post-task 4	DMb Post-task 4	Na Post-task 5	DMb Post-task 5	Na	DMb
cycle 2	DMa Post-task 1	.Nb Post-task 1	DMa Post-task 2	Nb Post-task 2	DMa Post-task 3	Nb Post-task 3	DMa Post-task 4	Nb Post-task 4	DMa Post-task 5	Nb Post-task 5	DMa	Nb
cycle 3	DMb Post-task 1	Na Post-task 1	DMb Post-task 2	Na Post-task 2	DMb Post-task 3	Na Post-task 3	DMb Post-task 4	Na Post-task 4	DMb Post-task 5	Na Post-task 5	DMb	Na
cycle 4	Nb Post-task 1	DMa Post-task 1	Nb Post-task 2	DMa Post-task 2	Nb Post-task 3	DMa Post-task 3	Nb Post-task 4	DMa Post-task 4	Nb Post-task 5	DMa Post-task 5	Nb	DMa

Note: 1. Na refers to the cartoon episode "puppy tale", Nb refers to "baby butch". DMa refers to the topic "cyber love", DMb refers to "unbalanced status".

2. For each group, the content of the post-task activity was indicated by the name of that group.

3.5 Setting

The study was conducted in a large university in south China. Permission was granted by that university. The participants attended the study voluntarily. Prior to the study, they were told that they would be recorded while they were doing the tasks, and the recordings would be confidential and anonymous which would be used only for research reasons. As rewards for their participation, when the study has been accomplished, their own digitalized recordings with an analysis report for their own performance were offered if they were curious abut that. All the participants were willing to join the study and contribute their recordings. No additional personal information was collected, except the major and the name of the participants for data identification. In the study, to protect participants' identities all the names are pseudonyms.

The ideal setting for the present would be a language multimedia classroom which is equipped with MP3 players for recording, and computers with Sound-scriber software for participants transcribing and word processing. However, these conditions were not available for the present study. In order to obtain an easy and convenient method for the transcribing activity, the researcher adopted cassette tape recorders for recording. The study took place in the classrooms where the students had regular classes. The researcher was the study teacher for the five meetings. Although the participants did not know the researcher prior to the study, they enjoyed all the sessions which was reflected in

the final informal interview. Participants transcribed their recordings from the tapes with paper and pencil. Afterwards, all the recordings were digitalized for research data transcribing, coding and analysis on the computer, and all the transcripts were collected by the researcher.

3.6 Procedures

The participants were visited five times at one week interval, the first time for orientation and the other four times for main study procedures.

Prior to the main study procedures, an orientation session was warranted to ensure that the participants were well informed of the task procedure and the basic transcribing skills (for the experimental groups). In addition, sample task practices were expected to reduce the participants' performance anxiety and free up their cognitive load to tasks in the main study (Bygate, 2000, 2001; Gass et al, 1999). As such, the first mission of the orientation session was task and transcribing training, and sample task practice.

As for the instruction on narrative tasks, not much guidance was needed since participants had done this kind of task before in their regular English lessons. The researcher gave certain instructions on how to retell a story after watching the video. A series of retelling strategies were explained in terms of story comprehension, discourse organization, and the selection of the language (e.g. the coherent use of tenses). Next the sample narrative task practice followed. An

episode named *Flying Cat* in *Tom and Jerry* was played to the participants. After two minutes planning, they were asked to describe the story to the recorder as if they were telling the story to someone else who had not watched the cartoon.

The next task was the decision-making task. Participants had done some similar discussion tasks in their regular lessons. However, they had not encountered this type of problem solving task. The researcher provided them with some suggestions on how to accomplish this type of task, and guided them to follow the steps as: a) to think of different sorts of advice which are possible for the situation; b) to discuss the advantages (the benefits and the feasibility of a certain advice) and the disadvantages (the difficulties or dangers of some advice); c) to compare those benefits and disadvantages so as to work out the best advice that could be offered. Then, participants were given a letter concerning the future of a boy who was excellent at both playing badminton and doing physics, and after two minutes planning they were asked to find out a best advice on the boy's future.

As far as the transcribing activity is concerned, only the experimental groups needed some training. The transcribing task was not strange to them, since there were frequent dictation exercises in their middle school English lessons. Dictation, which has been adopted in ESL as a long tradition, features complete or partial transcribing (Lynch, 2001; Swain, 1995). Participants were informed that they were going to transcribe their own performance recordings in the study. The basic

procedure and skills was reviewed for them. Given the many dictation exercises they had had in their middle schools, no transcribing practice was offered during the orientation period.

As for the revising groups, they were instructed to revise their transcripts mainly in the following two ways: a) to correct any mistakes in the original transcripts in terms of spelling, lexical errors, morphological and grammatical errors, collocation problems and content misunderstandings); b) to elaborate the transcripts by adding something which they should have said or by replacing some expressions with better options. No dictionaries or other references could be used at revising stage. Like the transcribing training, no specific revising practice was conducted.

In addition to the above task training, the second purpose for the orientation session was to find whether the research groups were comparable with regard to their English proficiency (as described in Section 3.2).

In the main study, the six groups were visited four times at weekly interval. Under each group, there were two sub-groups which did the same tasks but in reverse order. Each week the experimental groups performed a task (either narrative or decision-making task), and then conducted post-task transcribing. In narrative tasks, they watched a cartoon video from *Tom and Jerry*, planned for two minutes, and described the story to the recorder. With decision making task, they were given a problem letter, planned for two minutes, and discussed

appropriate advice in pairs or in small groups[1]. No time limits were set for either the task performance or the post-task activity. Different episodes and topics were arranged in a balanced order. Table 3.7 shows the arrangement of the cartoon episodes and topics. This arrangement was applied to all the subgroups in the study.

Table 3.7. Arrangement of Cartoon Episodes and Discussion Topics

	Sub-group 1	Sub-group 2
Cycle 1	N: *Puppy Tale*	DM: *Unbalanced Degree*
Cycle 2	DM: *Cyber Love*	N: *Baby Butch*
Cycle 3	DM: *Unbalanced Degree*	N: *Puppy Tale*
Cycle 4	N: *Baby Butch*	DM: *Cyber Love*

(Note: N stands for the narrative task, while DM stands for the decision-making task)

Next, different groups were assigned different post-task transcribing activities. The five transcribing conditions were operationalized as follows:

1. In the individual transcribing only group, after task performance, the participants were asked to transcribe part of their own performances.

2. In the individual transcribing and revising group, after task performance, the participants were asked to first transcribe part of their own performances, and then revise the original transcripts by themselves individually.

3. In the pair transcribing group, after task performance, the participants were asked to transcribe in pairs each member's extracted performance.

4. In the pair transcribing and revising group, the participants were asked first to transcribe in pairs each member's extracted performance, and then revised the two transcripts in pairs.

5. In the pair transcribing and teacher revising group, the participants were asked to transcribe in pairs each member's extracted performance. Then, their transcripts were revised by the teacher (the researcher of the study). At the beginning of the following task cycle, the teacher-revised transcripts were returned to the participants for them to read.

Considering the length of the extracted performance, in a narrative task, the participants who worked individually transcribed a 3-minute performance starting from a certain story point. For the pair work groups, things were more complicated. Each member of the dyad contributed a 1.5-minute performance in a continuous storyline. For example, starting from a story point (*The door is too small for the puppy to come in*...), participant A's 1.5-minute performance were transcribed, and the transcription ended with another story point (*Tom was angry*...) from which participant B's 1.5-minute performance started. In this way, the story content contributed by the dyads was supposed to be different from each other so as to avoid any competition concerning the quality of the oral performance between the two members. Several story turning points in the cartoons were provided for free selection by each pair. In a decision making task, considering the time for turn taking, for between-interlocutors' pauses which

tended to occur more frequently than that in the narrative task, five-minute performance was assigned to be transcribed either individually or in pairs.

As for the control group, they did the same tasks as the experimental groups did. However, no post task activities were involved after the task performance.

At the end of the last task cycle, an informal interview was carried out to seek for some feedback on the study, specifically participants' reflection on their task performance and on the post-task activities.

3.7 Data transcription and coding

As mentioned above, the original data were recorded by cassette tape recorders for participants transcribing. Afterwards, the data were digitalized into computer-readable wave files for research transcribing, coding and analysis.

The data were first transcribed into word documents by using Sound Scriber software. Then, the transcribed speech was further divided into the analysis of speech units (AS-unit) for computerized analyses to assess oral performance features. The following are the coding procedures for the present study.

First of all, the guidelines of Foster et al. (2000) were followed closely in dividing the discourse into AS-units. Independent clauses and subordinate clauses, as well as independent sub-clausal units which could be elaborated into full clauses, were counted as AS-units. In certain more difficult cases, pauses and intonation were taken into consideration as the governing principle was whether

or not a piece of discourse constituted a single chunk of micro-planning. The following two examples illustrate some difficult cases, such as coordinated verb phrases. For this case, the combined attention to both mid-clause pause length and the intonation contours is clearly warranted.

Example 1: |*Puppy open the door and call his mate to share the milk*| (1 AS-unit) (pause between coordinated phrases is less than 0.5 seconds.)

Example 2: |*Jerry helped Tom into the house* (0.6)| *and give some hot water to Tom*| (2 AS-unit) (the first coordinated phrase is marked by falling intonation and pause between two coordinated phrases is longer than 0.5 seconds)

Subsequently, the documents with AS-unit divisions were further transferred according to two formats: one format is CHAT format which can be readable for *vocd* program in the CHILDES project to measure the lexical diversity of the data (see Section 2.5 for task measures); the other format is Task Profile (TP) format which is readable by the *Task Profile* program, written by Skehan, to assess the complexity, accuracy and fluency, and also the lexical variation. Based on the AS-unit segmentation, the CHAT-formatted and TP-formatted transcripts were further coded for measures of complexity, accuracy and fluency.

As far as complexity is concerned, the ratio of clauses to AS-units and the mean length of AS-units (i.e. the number of words per AS-unit) are the complexity indicators. Clause coding also closely followed Foster et al. (2000)'s guidelines. Generally, non-finite verbs were only coded as clauses when they were

accompanied by a subject, object, complement or adverbial to verify their clausal status. The number of clauses produced by each participant was divided by the total number of AS-units to arrive at the complexity ratios. The clauses in AS-units are coded in the following way:

Example 3: |*they decide to eat*| (1 clauses, 1 AS-unit, no clausal element following the non-finite verb)

Example 4: |*Um He run :: to get the ham*| (2 clauses, 1 As-unit)

Example 5: |*Tom know :: that there was a dog in the house*| (2 clauses, 1 AS-unit)

With regard to the separate dimensions of fluency, different codes were adopted to make the features readable by computer program. Repair fluency was coded as the following:

1) false starts:

Example 6: |{*He he he um he he get*}# *the cat think an idea*|

Example 7: |{*and Jerry didn't*} # *and Jerry came*|

2) reformulations:

Example 8: | *he saw a strange thing* {*which are*}~ *which was heavy*|

Example 9: |{*He found that the thing*}~ *he found out that the thing was seven dogs*|

3) replacements:

Example 10: | {*He want to pick*} rpl *he want to get the milk* |

Example 11: | *and {the bags}*rpl *the dogs in the bags came out*|

4) repetitions:

Example 12: | *{and}* and he {threw}* threw them to the house*|

Example 13: |*Um Jerry {went across}* went across and {find the bag}* find the bag*|

The timing issue is crucially significant in the calculation of mean length of run, and speech rate. As such, the following timing indicators were marked: the starting and ending time of the whole speech, of each AS-unit.

Examples 14:

|And then the five dogs left.|

<00.20.26><00.25.03> (the starting and ending time of the AS-unit)

As for accuracy, the identification of errors reflects the number and the ratio of error-free clauses. Error-free clauses were defined as clause in which no error was made in terms of syntax, phonology, morphology, native-like lexical choice and discourse (Skehan, coding manual). However, errors in stress, intonation, patterns or pronunciation of the words and utterances were not included. The following are the examples of accuracy coding:

Example 15: |*It's a dark night.* errfr |

Example 16: |*The little dog want to* err_m_m :: *stay with Jerry.* errfr | (to indicate separately whether each clause in an AS-unit is error free or erroneous.)

In the current study, all the data, divided into AS-units, were transcribed in

both CHAT and TP formats by using the above coding marks for quantitative analysis on participants' oral performance. In addition, two procedures were adopted to ensure the reliability of data coding: intra-coder and inter-coder checks. The researcher recoded a subset of the data (30% of the data set) for a second time with an interval of four weeks after the first coding. The agreement of the intra-coding was 97%. As for the inter-coder reliability, a Ph.D candidate who also did the task-based research and coded his own data according to the same coding manual (Skehan, coding manual) was invited to code a subset of the data (10% of the data set). The agreement of the inter-coding was 95%. The agreements in both the intra-coding and the inter-coding procedures established the reliability of the data coding in the present study.

3.8 Analyses methods

Data were analysed using the statistical package SPSS. To address research question one concerning the effects of post-task transcribing, a one-way Multivariate Analysis of Variance (MANOVA) was performed.

For research questions 2 and 3 concerning the effects of pair/individual transcribing and the effects of further revision, two-way MANOVAs were performed to consider the two independent variables simultaneously.

For research question 4 which attempts to explore the different effects of three revision conditions, that is the individual-based revision, pair-based revision

and teacher-involved revision, a one-way MANOVA was used.

All the MANOVAs were followed by post-hoc comparisons of all the examined conditions to identify which groups were significantly different from the others.

For research questions 5 and 6, which explore the within-subject effects, namely the effect of task types and task practices, repeated measures MANOVAs were performed to see the within-subject effects of both task types and task practices. Then, post-hoc comparisons were conducted for the significant effects to identify the significantly different groups.

Chapter Four Data Analysis on Post-task Transcribing Effects on Narrative Task Performance

4.0 Introduction

There are two major concerns of the present research—the effects of post-task activities and of task practice. This chapter and the next one will focus on the effects of post-task transcribing in narrative and interactive tasks respectively.

The presentation of the results concerning the narrative task performance is organized to address research questions one to four in sequence. In section 4.1, for research question 1, the focus is on the general effects of post-task transcribing by comparing all the five post-task groups with the no post-task control group. In section 4.2, to address research questions 2 and 3, the foci are on the effects of individual/pair conditions and the effect of revision on learners' task performance. In section 4.3, for research question 4, the focus is on the differences among the three revision conditions.

In the last section (section 4.4), a synthesis of the findings is provided to show a complete picture of the results regarding the effects of post-task transcribing in narrative task performance.

4.1 Data Analysis regarding Research Question 1

Research question 1 aims to explore whether there is any effect of post-task transcribing as compared to the non-post-task condition (i.e. the control group) on learners' task performance. As the major concern of the current research, the effect of post-task transcribing is investigated by considering performance in the two narrative tasks: the first narrative task performance will be viewed as the baseline for further comparison, and the second narrative task performance will be investigated to see whether the post-task treatment groups, as a result of the involvement of post-task transcribing, are significantly different from the control group.

4.1.1 Fluency

This section concerns the effects of post-task transcribing on fluency. First, the descriptive statistics for narrative fluency are presented in Table 4.1. The results show that as compared to the five treatment groups, in the first narrative task performance the control group used the least replacements and false starts, but most reformulations, whereas in the second narrative task performance the control group performed at the lowest speech rate and used most repetitions. The data are normally distributed with no excessive kurtorsis or skewedness.

Table 4.1. Descriptive Statistics on Narrative Task Performance: Fluency

Groups (n=16 for each)	filled pauses				words_per_minute				reformulation				repetition				replacement				false starts			
	N1		N2		N1		N2		N1		N2		N1		N2		N1		N2		N1		N2	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
individual	5.94	4 54	4.44	3.52	59.82	10 74	74 69	23.01	1 25	1.29	1.31	1 49	16.37	10 34	16.37	11 09	2 25	2 38	2.31	1 7	56	1 15	2 25	2.38
individual revise	5 13	3 34	6 94	4.48	70 00	26 86	71 28	17 06	1 25	1 06	2 00	2 19	12 06	11 21	14 93	10 24	2 06	1 69	2 50	1 54	1 06	1 28	2.06	1 69
pair	5 37	3 67	4 51	2.95	54 61	15 91	67 20	8 73	93	85	1 18	1 16	18 62	13.59	11 81	6 16	2 43	1 96	2 93	1 18	1 62	1 68	2 43	1 96
pair revise	6 00	5 63	7 25	6.15	67 58	12 04	74 71	12 13	1 81	1 32	1 81	1.33	16 31	8.54	16 87	9 07	2 00	1 31	1 68	1 3	1 12	80	2 0	1 31
pair teacher revise	6 94	3 66	7 44	3.20	71 33	17 55	75 43	19 06	1 31	1 07	1 81	1 60	13 56	8.29	14 62	6 32	1 93	99	2.06	1 4	1 18	91	1 93	99
control	5 81	6 77	4.87	5 19	65 17	14 53	66 89	10.54	2 00	1 21	1 93	1 48	16.75	10 45	17 31	12 62	1 81	1 51	1 81	2 42	62	61	1 81	1 51

To explore the statistical differences between the post-task groups and the control group, one-way MANOVAs were performed for the first and second narrative task performances respectively. MANOVA results for the first narrative task performance show that there is no significant difference among the six groups in terms of fluency in the first narrative task performance. In other words, in the first narrative task performance, the six groups were indistinguishable as far as fluency is concerned.

For the second narrative task performance, the MANOVA results show that there is no significant difference among the six groups in terms of fluency as well.

As a result, it can be said that even if the post-task treatment groups have experienced post-task transcribing, they still did not outperform the control group with regard to fluency which means that post-task transcribing has no significant effect on learners' oral fluency.

4.1.2 Complexity

Post-task transcribing, as a focus-on-form activity, is expected to be beneficial for the formal aspects of task performance. The formal aspects include complexity, accuracy and lexical performance in the present study. This section focuses on the effects of post-task transcribing on complexity in narrative tasks. The descriptive statistics (see Table 4.2) show that in comparison with the five treatment groups, in the first narrative task, the control group employed fewest words per AS-unit, whereas in the second narrative task, the control group used the fewest clauses and words per

AS-unit.

Table 4.2. Descriptive Statistics on Narrative Task Performance: complexity

Groups (n=16 for each)	Clauses per AS-unit				Words per AS-unit			
	N1		N2		N1		N2	
	mean	SD	mean	SD	mean	SD	mean	SD
individual	1.25	.12	1.38	.13	7.84	.86	9.20	1.25
individual revise	1.33	.12	1.42	.15	8.49	1.75	9.44	1.33
pair	1.29	.16	1.53	.13	7.95	1.34	8.79	1.29
pair revise	1.30	.12	1.47	.14	7.79	1.06	8.16	1.30
pair teacher revise	1.33	.10	1.42	.14	7.74	1.85	8.47	1.33
control	1.32	.14	1.35	.12	7.60	1.31	7.80	1.32

To see the significant differences among the groups, one-way MANOVAs were carried out. For the first narrative task, the MANOVA results show that there is no significant difference among all the six groups: F (20,289) =.97, p= .499.

Another MANOVA was performed on the second narrative task performance, and the results (see Table 4.3) show that there is a significant difference among the groups: F (10,178)=2.79, p=.003. Subsequent univariate tests examined the two complexity measures separately. A significant difference was reported among the groups (p=.007) with regard to the number of clauses per AS-unit, and a near significance concerning the mean length of AS-unit (i.e. the number of words per AS-unit) (p=.056).

Table 4.3. Univariate Tests on the Second Narrative Task Performance: Complexity

source	dependent variables	df	mean square	F	sig.
groups	clauses per AS-unit N2	5	.062	3.43	.007*
	words_per_AS-unit N2	5	6.146	2.25	.056

(* p<.05)

In view of this significant result, Post hoc analyses (Tukey HSD) were performed to identify which post-task transcribing groups significantly differed from the control group. The results show that when the number of clauses per AS-unit is concerned, only the pair transcribing group is significantly better than the control group. Effect size calculation was conducted to see the magnitude of the significant effect, and Cohen's *d* [5]value was 1.43 which can be treated as a very large effect.

In short, as for the effects of the post-task transcribing on complexity, whereas no significant differences were found in the first narrative task performance, a significant difference emerged between the control group and one post-task transcribing group (the pair transcribing group) in the second narrative task. In particular, the pair transcribing group produced significantly more clauses than the control group. Since only one (out of five) post-task transcribing group significantly outperformed the control group, it may only suggest a limited (or at least, a focused) effect of post-task transcribing on complexity in the narrative task performance.

[5] Cohen (1988) defined effect sizes as "small, d< 0.4" "medium, 0.4<d<0.75", and "large, 0.75<d<1.10", "very large, 1.10<d<1.45", "huge, 1.45<d".

4.1.3 Accuracy

In addition to complexity, accuracy is another important (or even the most important) aspect to be examined with respect to the effects of a focus on form activity. This section examined the effects of post-task transcribing on accuracy. The descriptive results (see Table 4.4) show that as compared with the post-task treatment groups, in the first narrative task performance, the control group was towards the middle-level when all the accuracy measures were concerned. In the second narrative task, the control group got the lowest error-free ratio and the shortest accurate clauses, and at the same time, the control group produced the most errors per 100 words.

Table 4.4. Descriptive Statistics on Narrative Task Performance: accuracy

	error-free ratio				errors per 100 words				accuracy length			
	N1		N2		N1		N2		N1		N2	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
individual	.39	.12	52	.13	12 46	5.88	9 61	3.94	5 57	2.99	4 81	2 69
individual revise	.37	.13	54	.11	13.36	5 60	8 87	3.26	4 19	2.76	6 5	2 36
pair	.34	11	.51	.10	14 71	4.34	10.35	3.47	2 87	1.96	5.06	1 84
pair revise	.32	.11	.51	.13	12 63	5 10	8.85	3 18	3.69	2 06	5 38	2.22
pair teacher revise	.37	.08	52	.12	11.32	5.72	10.08	3 93	3 19	1 11	5 69	2 24
control	.35	.11	.38	.08	13 14	5.20	12.06	5.49	3 06	2 35	3 94	2.57

MANOVA results for the first narrative task performance reveal that in the first narrative task performance there is no significant difference among all the groups regarding accuracy. In contrast, MANOVA results for the second narrative task performance show that the groups were significant different in the second narrative

task regarding accuracy: F (15, 243) =1.794, p=.036. Subsequent univariate tests show (see Table 4.5) that a significant difference exists among the groups regarding the proportion of error free clauses (p=.001), and an approaching significance with regard to accuracy length (p=.064), but no significance on errors per 100 words.

Table 4.5. Univariate Tests on the Second Narrative Task Performance: Accuracy

Source	dependent variables	mean square	F	sig.
Groups	ratio of error free clauses_N2	.057	4.384	.001*
	errors_per_100 words_N2	22.69	1.45	.214
	accuracy length _N2	11.89	2.18	.064

(*p<.005)

Post-hoc analyses (Tukey HSD) show that all five post-task transcribing groups are significantly better than the control group in terms of error-free ratio. In view of the significant results, effect sizes were calculated. The significances and their effect sizes (Cohen's d) are presented in Table 4.6.

Table 4.6. Significances and Effect Sizes for Accuracy in the Second Narrative Task Performance

	treatment groups	sig	effect sizes & its level
control <	individual transcribing	.012*	1.34, very large
control <	individual transcribing and revising	.001*	1.72, huge
control <	pair transcribing	.019*	1.48, huge
control <	pair transcribing and revising	.014*	1.24, very large
control <	pair transcribing and teacher revising	.009*	1.42, very large

(*p<.05)

Thanks to the involvement of the post-task transcribing, all the post-task groups were significantly better than the control group in the second narrative task when the

proportion of error-free clauses is concerned. The adoption of transcribing at the post-task stage, therefore, proved to be effective to promote learners' accuracy, in particular to encourage them to produce more error-free clauses. In addition, a near significance was reported with regard to the accuracy length which may further suggest a trend supporting the post-task transcribing groups to use longer accurate clauses in the narrative task performance than the control group did.

4.1.4 Lexical Performance

Lexical performance, independent of the above syntactic aspects (i.e. complexity & accuracy), is another indispensable concern in exploring the effects of post-task focus on form activities. In this section, we will compare the lexical performance of post-task transcribing groups to that of the control group. Two lexical aspects are examined: lexical sophistication and lexical diversity. The descriptive statistics for the lexical performances are presented in Table 4.7. In the first narrative task, the control group was towards the middle level concerning the mean scores on lexical sophistication, but they got the lowest scores on lexical diversity. In the second narrative task, the mean score of lexical sophistication of the control group is the lowest, whereas its mean score of lexical diversity is the highest among the six groups.

Table 4.7. Descriptive Statistics on Lexical Performance in Narrative Tasks

Groups	lexical diversity				lexical sophistication			
(n=16 for each)	N1		N2		N1		N2	
	mean	SD	mean	SD	mean	SD	mean	SD
Individual	30.75	5.33	36.78	6.28	1.99	.37	2.12	.33
individual revise	34.65	13.09	33.28	5.97	1.89	.30	2.43	.11
pair	29.54	6.96	34.43	5.49	1.86	.39	2.17	.28
pair revise	31.44	5.63	35.78	5.09	1.89	.41	2.11	.30
pair teacher revise	29.23	7.03	34.72	6.45	1.94	.35	2.20	.19
control	29.55	6.58	37.78	7.69	1.93	.56	2.04	.48

As for the first narrative task performance, MANOVA results show no significant differences among the groups in terms of lexical performance.

Another one-way MANOVA was performed on the second narrative task performance. It reveals a significant difference among the groups in terms of lexical performance: $F(10,178)=2.242$, $p=.017$. In the subsequent univariate tests, the significance is reported in terms of lexical sophistication, but not on lexical diversity. The results are presented in Table 4.8. Post hoc analyses (Tukey HSD) were carried out and showed that as far as the effect of post-task transcribing on lexical sophistication was concerned, only the individual transcribing plus revision group produced significantly more infrequent words than the control group ($p=.007$), and the effect size for the significance is 1.16 which is treated as a very large effect. No other post-task treatment groups demonstrated the effects of post-task transcribing on lexical performance.

Table 4.8. Univariate Tests on the Lexical Performance in the Second Narrative Task

Source	dependent variables	mean square	F	sig.
Groups	lexical diversity N2	43.538	1.126	.35
	lexical sophistication N2	.287	3.111	.01*

(* p<.05)

Given that only one treatment group is significantly superior to the control group in terms of lexical sophistication, the effect of post-task transcribing may only be considered as a partial effect on lexical aspects of narrative task performance.

4.1.5 Summary

The first research question explores that whether post-task transcribing has effects on different aspects of learners' oral performance. To address this question, a series of comparisons were carried out between the post-task transcribing groups and the control group in terms of different aspects of narrative task performance, i.e. fluency, complexity, accuracy and lexical performance.

In the first narrative task performance, no significant differences were reported between the post-task groups and the control group with regard to all the concerned aspects of task performance. It is, therefore, argued that the six groups were indistinguishable when they were doing the first narrative task and were comparable for further treatment.

In exploring the effects of a post-task transcribing activity, the second narrative task performance is the lens in the analyses. With regard to fluency, none of the post-task transcribing group was significantly superior to the control group which suggested that being involved in different types of post-task transcribing activities did not bring about any significant effects on learners' fluency.

As for complexity, one (out of five) post-task transcribing group (i.e. the pair transcribing group) produced significantly more clauses than the control group. Since the effect on complexity is limited to one group, and does not emerge in most post-task transcribing groups, we may only suggest a selective effect of post-task transcribing on complexity in narrative tasks. Similarly, lexical performance is another aspect in which a partial effect of post-task transcribing is reported. Only the individual-transcribing-and revision group outperformed the control group by employing more infrequent words in the narrative tasks, whereas no such effect is found among the other post-task groups or regarding lexical diversity.

As compared to the above measures which demonstrate no effect or limited effects of post-task transcribing, accuracy proves to be the most encouraging aspect to show the consistent effects of post-task transcribing. All the post-task transcribing groups outperformed the control group significantly when the proportion of error-free clauses is concerned. This suggests that post-task transcribing activities, although conducted in various conditions, are consistently effective in pushing learners to focus on form, and consequently to improve their language accuracy.

In sum, post-task transcribing, adopted as a focus on form activity in the current research, has a clear and consistent effect on accuracy, partial and selective effects on complexity and lexical performance, but no effects on fluency in narrative task performance. To explore the effects of focus on form activity in task-based language learning, the findings in narrative task performance are encouraging. A brief interpretation of the results is presented in section 4.4, and detailed discussion on the findings will be presented in Chapter 7.

In the above, we explored the general effects of post-task transcribing on learners' task performance. Furthermore, in this research, the post-task transcribing is operated under different conditions. In the following two sections (section 4.2 and 4.3), the foci will be on the different effects of the various post-task transcribing conditions. Since no post-task transcribing has been employed in the control group, only the post-task transcribing groups are included in the analyses. In addition, although all the participants were involved in two narrative tasks, the counterbalanced design of the study (cf. chapter 3) determined that some subgroups did not experience any post-task activities prior to the first narrative task, while other subgroups did. However, in the second narrative task, all the participants have been involved in post-task transcribing. Thus, in the following study, only the second narrative task performance is examined.

4.2 Data Analysis regarding Research Question 2 and 3

Research questions 2 and 3 both pertain to the effects of different factors in

post-task transcribing. In particular, research question 2 considers whether there are any different effects between the individual-based and the pair-based post-task transcribing, and research question 3 aims to investigate whether there is any effect of the involvement of further revision after post-task transcribing as compared to the non-revision condition.

The two research questions are related to each other in that the two independent variables (i.e. the individual/pair condition and the involvement of revision) co-occur among the post-task groups. If these two independent variables are examined separately, one of them will possibly be the intervening factor in the analysis of the other. Furthermore, it is also possible that an interaction effect of the individual/pair condition and of the involvement of revision occurs among the groups. As a result, it is necessary to consider the effects of both factors together, rather than separately, by performing two-way MANOVAs on the second narrative task performance to see the effects of the individual/pair condition, of further revision, and of the interaction effects as well. In the following analyses, four post-task treatment groups will be examined with the exclusion of the pair-transcribing-teacher-revision group so as to remove the intervening factor of the teacher's role in that group: The role of the teacher will be investigated in section 4.3. As in section 4.1, the analyses are presented around the different aspects of task performance in four sub-sections. The findings will be summarized at the end.

4.2.1 Fluency

This section focuses on the effects of the individual/pair transcribing and the effect of revision on fluency. The descriptive statistics are given in Table 4.9. With regard to the differences between the individual and the pair condition, the descriptive statistics show that individual transcribing groups produced more reformulations, repetitions and replacements than the pair transcribing groups, but fewer filled pauses and false starts than the latter, whereas the two conditions produced similar numbers of words per minute. As for the comparison between the revision and the no revision conditions, the descriptive statistics reveal that except for the employment of replacements, the revision condition produced more fluency indices than the non-revision condition in terms of filled pauses, words per minute, reformulation, repetition and false starts.

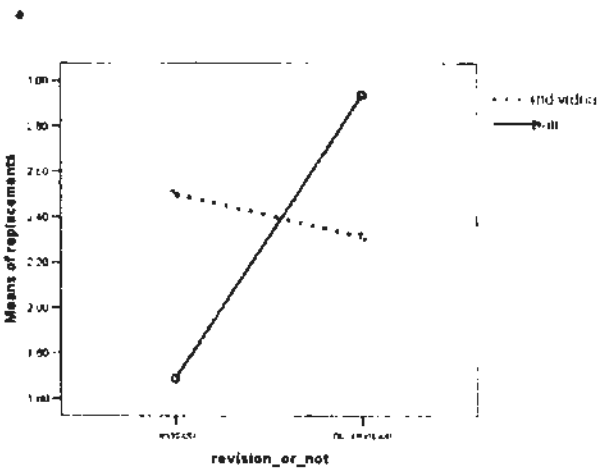
Table 4.9. Descriptive Statistics of the four treatment groups in the narrative task performance: fluency

Conditions (n=16 for each group)	filled pauses		words per min		reformulations		repetitions		replacements		false starts	
	indivi dual	pair	indivi dual	pair	indivi dual	pair	individu al	pair	indivi dual	pair	individu al	pair
	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)
revision	6.93 (4.48)	7.25 (6.15)	71.28 (17.1)	74.71 (12.1)	2.00 (2.09)	1.81 (1.33)	14.94 (10.25)	16.88 (9.08)	2.50 (1.55)	1.69 (1.30)	1.06 (1.29)	1.13 (.81)
no-revision	4.44 (3.52)	4.50 (2.90)	74.70 (23.0)	67.21 (8.74)	1.31 (1.49)	1.19 (1.17)	16.38 (11.10)	11.81 (6.17)	2.31 (1.70)	2.94 (1.18)	.57 (1.15)	1.63 (1.67)

A two-way MANOVA was performed and significance is noted for the interaction effect between the individual/pair condition and revision: $F(6,71)=2.39$, $p=.044$. But

there are no significant main effects in terms of either the individual/pair contrast (p=.62), or the involvement of revision (p=.11). Univariate tests reveal that the interaction effect is demonstrated on the adoption of replacements (p=.05). Examination of the plot (see Figure 4.1) suggested that when the post-task transcribing was performed individually with further revision, more adoption of replacements emerged than the pair transcribing with revision. On the other hand, when the post-task transcribing was conducted in pairs without further revision, the employment of replacements was more frequent than the individual transcribing without revision.

Figure 4.1 Means of Replacements in the Narrative Task



In terms of fluency, an interaction effect between the individual/pair condition and revision emerged on one fluency measure: replacements. That is to say, the effect of either individual or pair condition on the production of replacements depends on the presence or absence of further revision. However, given that the interaction effect is limited to only one (out of six) fluency measures, it is premature to suggest that the interaction of both factors (the individual/pair transcribing and the involvement of

revision) has significant effects on fluency in narrative task performance.

4.2.2. Complexity

This section concerns the effects of the individual/pair condition and the effects of revision on complexity. The descriptive statistics are shown in Table 4.10. The descriptive results show that on the one hand, the pair transcribing groups produced a larger number of clauses per AS-unit than the individual-based groups, but used fewer words per AS-unit than the latter. On the other hand, the revision condition produced fewer clauses and words per AS-unit than the no revision condition.

Table 4.10. Descriptive Statistics of the four treatment groups: complexity

conditions	clauses per AS-unit		words per AS-unit	
	individual	pair	individual	pair
	mean (SD)	mean (SD)	mean (SD)	mean (SD)
revision	1.43 (.15)	1.47 (.14)	9.44 (2.10)	8.17 (1.81)
no revision	1.38 (.13)	1.53 (.13)	9.20 (1.18)	8.79 (1.54)

A two-way MANOVA showed that there was no significant interaction of the two factors. As for the main effects of each factor, whereas no significance was noted with regard to the effects of revision, the individual and pair-based transcribing groups were significantly different regarding complexity: $F(2,59)= 6.87$, $p=.002$. Univariate tests showed that the significant differences emerged in terms of both the number of clauses per AS-unit ($p=.007$); and of the number of words per AS-unit ($p=.050$). Effect sizes (Cohen's *d* values) for the significance are medium. The results are presented in Table

4.11.

Table 4.11. Univariate tests of Individual/Pair Conditions: Complexity

groups	dependent variables	F	sig.	effect size
pair > individual	clauses per AS-unit	7.69	.007*	0.7, medium
individual > pair	words per AS-unit	3.96	.050*	0.51, medium

Despite the significance noted above, it is still not clear about the different effects of individual/pair conditions on complexity, because the results (see Table 4.11) show contrasting effects with regard to the two different complexity measures. In particular, the pair transcribing groups produced significantly more clauses per AS-unit than the individual-based groups. On the other hand, the latter used significantly more words per AS-unit than the former. It may, therefore, be inferred that the pair transcribing groups produced more but shorter clauses or clause elements than the individual condition; whereas the latter adopted more words but simple syntax (with less subordination) in the narrative tasks. The finding, different from the results in previous studies, is returned to in section 4.4 and further discussed in Chapter 7.

4.2.3 Accuracy

This section concerns the effect of the individual/pair condition and the effect of revision on accuracy. The descriptive statistics show (see Table 4.12) that as for the comparison between the individual- and the pair-based groups, the individual condition produced a slightly larger proportion of error-free clauses and lengthier accurate

clauses than the pair condition. However, the former produced slightly more errors per 100 words than the latter. As for the comparison between the revision and no revision condition, the revision groups produced more error-free clauses and fewer errors and lengthier clauses than the non revision groups.

Table 4.12. Descriptive Statistics of the Four Treatment Groups: Accuracy

conditions	ratio of error-free clauses		errors per 100 words		accuracy length	
	individual	pair	individual	pair	individual	pair
	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)
revision	.55 (.11)	.51 (.13)	8.88 (3.26)	8.85 (3.18)	6.50 (2.37)	5.38 (2.22)
no revision	.52 (.13)	.51 (.10)	9.61 (3.94)	10.36 (3.47)	4.81 (2.69)	5.06 (1.84)

A two-way MANOVA showed no interaction effect for the two factors in terms of accuracy. In addition, there is no significant main effect of ether the individual/pair condition or the involvement of revision. In other words, whether the post-task transcribing is performed individually or in pairs, or whether there is a further revision gives rise to no different effects on accuracy in narrative task performance.

4.2.4 Lexical Performance

This section explores the effects of individual- or pair-based post-task transcribing and the effects of further revision on lexical performance. The analyses are around two aspects of lexical performance: lexical sophistication and lexical diversity. The

descriptive statistics are presented in Table 4.13. For the comparison between individual and pair conditions, the individual post-task transcribing groups got higher mean scores regarding both lexical diversity and lexical sophistication. When the revision groups are compared to the no revision ones, the revision groups got a higher mean score for lexical sophistication, but a lower mean score for lexical diversity.

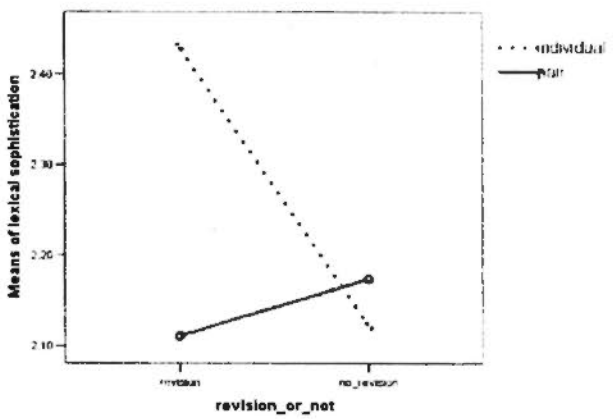
Table 4.13. Descriptive Statistics of the Four Treatment Groups: Lexical Performance

conditions	lexical diversity		lexical sophistication	
	individual	pair	individual	Pair
	mean (SD)	mean (SD)	mean (SD)	mean (SD)
revision	33.28 (5.97)	35.79 (5.10)	2.43 (.11)	2.11 (.30)
no revision	36.78 (6.28)	34.43 (5.49)	2.12 (.32)	2.17 (.28)

The results of a two-way MANOVA showed that first of all, there was a significant interaction effect of the individual/pair condition and the revision involvement: F (2.59)=5.34, p=.007. However, there are no significant main effects for both factors respectively. Further univariate tests show that between the two aspects of lexical performance, the significant interaction effect pertains to lexical sophistication (p=.007, Cohen's d=0.7, a medium effect), rather than to lexical diversity (p=.095). Inspection of Figure 4.2 suggests that both the individual and the pair conditions, with or without further revision, demonstrate distinct effects. In particular, when further revision is involved, the individual condition generates a better performance in terms of lexical sophistication (i.e. the use of more infrequent words) than the pair condition.

However, as long as no further revision is involved, the pair transcribing group produced more infrequent, sophisticated lexical items than the individual transcribing group.

Figure 4.2. Means of Lexical Sophistication in Narrative Tasks



In brief, with regard to lexical sophistication, an interaction effect of both the individual/pair transcribing and revision was reported. The presence of further revision supported the individual transcribing group to use more infrequent words, whereas the pair transcribing group took the advantage of no revision involvement and adopted more infrequent words in the narrative task performance. The interaction effect which is seldom researched with this respect in previous literature receives further attention in section 4.4 and more in Chapter 7.

4.2.5 Summary

Research question 2 concerns that whether the individual and the pair-based

post-task transcribing have different effects on task performance. Research question 3 aims to explore whether further revision after post-task transcribing has a significant effect on performance. Instead of addressing the two questions separately, we used a two-way MANOVA to explore a) the main effects of the two independent variables (i.e. the individual/pair condition, and the involvement of revision), b) the interaction effects of both factors on different aspects of task performance.

As far as fluency is concerned, neither the individual/pair condition nor the involvement of revision has main effects on fluency, although an interaction effect is found on the use of a repair fluency measure: replacements. Except for this significant fluency indicator, the other five fluency measures did not show any significance. We can, therefore, only suggest that generally speaking, either the individual/pair condition or the revision involvement has no clearly significant effects on fluency.

Complexity is one of the major concerns of the present research. When the two independent variables are examined, whereas no interaction effect and no main effect of revision are reported, the significant effect of the individual/pair condition is noted on both complexity measures. However, the results for the two complexity measures cannot be patterned together: while the pair condition promotes the adoption of subordination, the individual condition encourages the use of more words per AS-unit. It may be inferred that although both are connected to the same construct (i.e. complexity), the complexity measures adopted in this study may reflect distinctive aspects of that construct. This finding will be interpreted further in Section 4.4 and

Chapter 7.

The results regarding the effects of both factors on accuracy are somewhat discouraging in that no significant main effects or interaction effect are reported regarding all the adopted accuracy measures. This suggests that whether the post-task transcribing is performed individually or in pairs has no different effects on accuracy. Neither does the involvement of revision.

When the effects are considered on lexical performance, an interaction effect on lexical sophistication is found. In particular, if post-task transcribing is performed individually, then revision is supportive to promote the adoption of infrequent words; on the other hand, when post-task transcribing is carried out in dyads, the absence of revision encourages learners to use more infrequent words. It may be suggested that with regard to lexical performance, the involvement of revision is not beneficial for all the learning conditions: it is effective for the individual context, but is a hindrance in the pair condition. Detailed discussions are presented in Chapter 7.

4.3 Data Analysis regarding Research Question 4

Research question 4 is a further exploration on the differences among the different revision conditions. Three different revision conditions are employed in the current study: one is teacher-involved revision, and the other two are student-based revision (i.e. individual-based revision and pair-based revision). The three revision groups are considered in this section. To look for differences among the three revision conditions,

a series of one-way MANOVAs were performed on various aspects of task performances in the second narrative task, with a summary provided at the end of the section.

4.3.1 Fluency

This section explores the differences among the three revision conditions with regard to fluency. The descriptive statistics show that among the three groups the teacher revision group produced most words per minute and most filled pauses and false starts, the pair revision group used most repetitions, and the individual revision group employed most reformulations and replacements. The results are given in Table 4.14.

Table 4.14. Descriptive Statistics for the Three Revision Groups regarding Fluency in the Narrative Task

conditions (n=16 for each)	filled pauses		words per min		reformulation		repetition		replacement		false starts	
	N2		N2		N2		N2		N2		N2	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
individual revision	6 94	4.48	71 29	17 06	2 0	2 19	14 94	10 25	2 5	1 55	1 06	1 29
pair revision	7 25	6 15	74 71	12 13	1 81	1 32	16 88	9 08	1 69	1 30	1 13	81
teacher revision	7 44	3 20	75 44	19 06	1 81	1 60	14 63	6 32	2 06	1 48	1 19	91

One-way MANOVA results show that as far as fluency is concerned, no significance is reported among the three groups. This suggests that the different revision conditions bring about similar effects on fluency, and so will not be discussed

further here.

4.3.2 Complexity

This section considers the differences between the three revision conditions in terms of complexity. The descriptive statistics show (see Table 4.15) that the pair revision group produced the most clauses per AS-unit, whereas the individual revision group produced most words per AS-unit, and the teacher revision group took the second place for both measures.

Table 4.15. Descriptive Statistics for the Three Revision Groups regarding Complexity in the Narrative Task

Conditions (n=16 for each)	clauses per AS-unit		words per AS-unit	
	N2		N2	
	mean	SD	mean	SD
individual revision	1.42	.15	9.43	2.10
pair revision	1.47	.14	8.16	1.81
teacher revision	1.42	.14	8.47	1.84

One-way MANOVA results show that there are no significant differences among the three groups with regard to complexity which indicates that the three different revision conditions generate similar performance on complexity.

4.3.3 Accuracy

This section focuses on the effects of different revision conditions in terms of accuracy. The descriptive statistics show that the individual revision group produced most error-free clauses, and the lengthiest accurate clauses among the three groups, whereas the teacher revision group produced most errors per 100 words. The results are given in Table 4.16.

Table 4.16. Descriptive Statistics for the Three Revision Groups regarding Narrative Task Accuracy

Condition (n=16 for each group)	percentage of error-free clauses		errors per 100 words		accuracy length	
	N2		N2		N2	
	mean	SD	Mean	SD	mean	SD
individual revision	.55	.11	8.88	3.26	6.5	2.36
pair revision	.51	.13	8.84	3.18	5.37	2.26
teacher revision	.52	.12	10.08	3.93	5.69	2.24

One-way MANOVA results showed that the three revision groups did not differ significantly when accuracy was considered which suggests that the differences of three revision conditions do not lead to different performance regarding accuracy.

4.3.4. Lexical Performance

In this section, the difference of three revision groups is concerned with respect to the lexical aspects. The descriptive statistics show that individual revision group got the highest lexical sophistication score, but the lowest lexical diversity score, whereas the pair revision group gained the highest lexical diversity score, but the lowest lexical

sophisfication score, and the teacher revision group took the second place for both lexical measures. The results are presented in Table 4.17.

Table 4.17. Descriptive Statistics for the Three Revision Groups regarding Narrative Task Lexical Performance

Condition (n=16 for each group)	lexical sophistication		lexical diversity	
	N2		N2	
	mean	SD	mean	SD
individual revision	2.43	.11	33.28	5.98
pair revision	2.11	.30	35.79	5.10
teacher revision	2.20	.19	34.72	6.45

One-way MANOVA results show that there is a significant difference among the three revision groups: $F(4,88)=4.24$, $p=.003$. Further univariate tests reveal that when the two aspects of lexical performance are considered separately, although no significance is reported in terms of lexical diversity, there is a significant difference among the three revision groups with regard to lexical sophistication ($p=.000$). The results are given in Table 4.18.

Table 4.18. Univariate Tests among the Three Revision Groups regarding Narrative Task Lexical Performance

source	dependent variables	df	F	sig.
groups	lexical diversity	2	.737	.484
	lexical sophistication	2	9.39	.000*

(*p<.05)

Post hoc analyses (Tukey HSD) show that the individual revision group is significantly better than the pair revision ($p=.000$) and teacher revision groups

(p=.010), however, no significant difference is found between the pair revision group and the teacher revision group. Effect sizes (Cohen's d) are 1.43, 1.56 which are treated as very large or even huge effects.

In short, in terms of lexical sophistication, the individual revision group was significantly better than the pair revision and the teacher revision group. Given that no significance was reported when the pair revision and the teacher revision are compared, we may only suggest a partial role of students' revision on lexical sophistication, and at the same time, the three revision conditions have no significant differences on lexical diversity.

4.3.5 Summary

To address research question 4 which concerns the different effects of revision conditions, the findings are presented as follows: a) the different revision conditions in the present study do not lead to different performance in terms of fluency, complexity and accuracy, b) a deep concern of this research question is on the relative role of the teacher in revision as compared to the role of students in revision. The role of the teacher in revision does not show any advantage to that of the students. Furthermore, one of the students' revision group (i.e. the individual revision group) outperformed the teacher revision group significantly in terms of lexical sophistication with a huge effect size noted.

4.4 Summary and brief interpretations of findings

Post-task transcribing, in the present research, is adopted as a focus on form activity in the hope that it would be beneficial for learners' language development in task-based language teaching. This chapter concerns the effects of post-task transcribing in narrative task performance.

Research Question 1 considers the general effects of post-task transcribing as compared to the non-post-task condition (i.e. the control group). First of all, there is a significant post-task effect on language accuracy. Various post-task transcribing conditions consistently foster more accurate language than the control group does. In addition, limited effects of post-task transcribing were found in terms of syntactic complexity and lexical sophistication. Last, there is no significant post-task effect in terms of fluency and lexical diversity. Parts of the results are in line with the findings in Foster & Skehan (forthcoming) in that in the narrative task performance, the effects of post-task transcribing are demonstrated in terms of accuracy which is a crucial index to show the effect of focus on form. Further, the limited effects on complexity and lexical performance are more valuable as compared to the previous research results which show no effect of post-task activity on complexity in narrative tasks (Foster & Skehan, forthcoming; Skehan & Foster, 1997). Finally, it is not surprising to see no significant effects on fluency, because as a measure of meaning processing, fluency is not a major concern in the present focus on form research. Further, the trade-off between meaning (fluency) and form (complexity and accuracy) (Skehan &

Foster, 1996; Skehan, 1998) makes it reasonable that the *focus on form* activity does not have effects on the *meaning* side of task performance.

In addition to the general effects of post-task transcribing, the differences of real-time performance in various conditions in post-task transcribing are considered. Research question 2 explores the difference between individual-based and pair-based post-task transcribing. The only significance is noted with respect to complexity. However, surprisingly, the effects of either individual or pair conditions on the two aspects of complexity are in complete contrast with each other: the pair condition favors the adoption of subordination per AS-unit, however, the individual condition promotes the production of more words per AS-unit and more lexical sophistication.

It seems, therefore, that both the individual and the pair conditions in the present study bring about effects on complexity, but in different ways. Although Norris & Ortega (2007) propose to adopt multivariate measures for the multifaceted syntactic complexity, the findings reveal that the different measures of complexity are not necessarily positively correlated with one another in different L2 research contexts. In the present research the two complexity measures reflect distinctive aspects of syntactic complexity. Further discussion will be presented in Chapter 7.

Research question 3 focuses on the effects of revision after transcribing as compared to the no revision condition. To remove the intervening factors in the analyses (c.f. section 4.2), the two independent factors in RQs 2 and 3 (i.e. the effects of individual/pair condition and the effects of revision) are examined together.

Although no significant main effects of revision are reported on all the aspects of task performance, an interaction effect of the revision condition and the individual/pair condition on lexical sophistication emerged in the analyses: the revision condition encourages the individual-based group to employ less frequent lexical items, but impedes the pair-based group from doing so, and the pair-based group adopted more infrequent words in a non revision condition.

The findings reveal that in L2 language teaching, a seemingly favorable condition, such as the involvement of revision, cannot be a panacea for different contexts. An exploration on the interaction of different factors would be beneficial for curriculum design and pedagogical practice. Furthermore, the fact that no significant effects of revision are found in terms of accuracy and complexity confirms some previous findings on the role of revision which argue that there are not any supportive effects of revision for language development in terms of grammar (Truscott, 1996, 2005, 2007).

Research question 4, as a developed question based on question 3, aims to compare the different effects of three revision conditions. Whereas no significant differences are reported for fluency, complexity and accuracy, the only significance is noted in terms of lexical sophistication: the individual revision group outperformed the other two groups by employing significantly more infrequent words in narrative task performance. The findings, although in contrast to those which find favorable results for teacher revision (Lynch, 2001,2007), are in line with some studies regarding the role of revision, and the role of the teacher in revision (Hyland & Hyland, 2006; Ferris,

2003). Thus, it is thought-provoking to raise further questions on the effects of different conditions of revision and the role of teacher at the post-task stage in TBLT.

This chapter concerns the effect of post-task transcribing on narrative task performance. The findings are supportive for post-task transcribing to be an effective activity to promote more accurate language in the narrative tasks. In addition, partial effects on complexity and lexical sophistication are reported. Furthermore, when the various conditions of post-task transcribing are compared, a number of fresh but encouraging results are noted. All the findings will be discussed in detail in Chapter 7. In the following chapter, the data in the interactive tasks are analysed to explore the effects of post-task transcribing on task performance in a slightly different context.

Chapter Five Data Analysis on Post-task Transcribing Effects on Interactive Task Performance

5.0 Introduction

Two task types have been adopted in this study: narrative and interactive tasks. The previous chapter focused on the narrative task performance to explore the effects of post-task transcribing. In the present chapter, the interactive task performance becomes the focus of the analyses to address research questions one to four in terms of fluency, complexity, accuracy and lexical performance.

In section 5.1, for research question 1, the general effects of post-task transcribing on interactive task performance are investigated by means of a series of one-way MANOVAs. In section 5.2, the effects of various conditions in post-task transcribing are explored, in particular the difference between individual- and pair-based post-task transcribing (for research question 2) and the role of further revision after transcribing (for research question 3), and two-way MANOVAs are employed in the analyses on different aspects of task performance. In section 5.3, for research question 4, the comparisons among the three revision groups are carried out so as to look for the differences of different revision conditions in which one-way MANOVAs are performed. This chapter ends with a summary section, together with preliminary interpretations for all the findings in both narrative and interactive tasks.

5.1 Data Analysis regarding Research Question 1

As the major focus in the current study, research question 1 concerns the issue that whether the involvement of post-task transcribing has an impact on learners' task performance, in particular in interactive tasks in this chapter. Two interactive tasks were performed in separate sessions by all the participants, and both are considered for different objectives: the first interactive task performance is analysed as the benchmark for further comparison, while the second performance is examined in attempt to explore the achievement of the post-task groups as compared to that of the no post-task control group. The following sub-sections are related to the different aspects of the interactive task performance.

5.1.1 Fluency

First of all, the general effects of post-task transcribing are examined with regard to fluency. The descriptive statistics (see Table 5.1) show that in the first interactive task, the control group produced most reformulations, and at the same time had the greatest speech rate (words per minute) among all the groups. In the second interactive task, the control group took the second place among all the groups when the different fluency indices were considered separately. The data are normally distributed.

Table 5.1. Descriptive Statistics on Interactive Task Performance: Fluency

Groups	filled pauses				words_per_minute				reformulation				repetition				replacement				false starts			
(n=16 for	DM1		DM2		DM1		DM2		DM1		DM 2		DM 1		DM 2		DM 1		DM 2		DM 1		DM 2	
each)	mean	SD	Mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
individual	2 44	2 73	2 56	2 44	74.25	10 88	78 73	7 58	94	1 0	1 13	1 15	4 19	3 76	4 63	3 34	1 0	97	1 13	1 02	25	58	19	40
individual revise	2 13	1 54	2 63	1 15	66 94	12 75	67 99	9 35	75	93	94	68	4.88	3 40	4 31	1.54	1 38	1 15	1 19	98	44	63	31	48
pair	4 06	2 41	3 25	2 18	70 48	17 26	72 43	15 63	63	50	75	93	6 19	1 76	6 13	1 54	1 13	96	1 38	1 31	19	46	19	40
pair revise	3 50	2 48	3 56	1 90	63 52	9 44	71 98	16 81	44	51	56	63	5 31	1 74	5 44	1 59	63	62	81	75	31	48	25	45
pair teacher revise	5 56	6 50	4 25	3 45	67 54	18 54	68 70	13 35	44	51	50	52	5 88	2 83	5 19	1 60	1 00	97	1 19	83	31	48	69	48
control	3 50	3 05	3 68	2 96	81 28	7 89	76 16	8 43	1 31	1 08	94	99	4 13	2 53	5 56	2 50	1 25	1 18	1 19	91	31	48	31	48

For the first interactive task performance, the MANOVA results show that there are significant differences among the groups F(30,342)=1.782, p=.008, which are found in terms of words per minute (i.e. speech rate) and reformulations (see Table 5.2).

Table5.2 Univariate tests on the First Interactive Task Performance: Fluency

source	dependent variables	df	F	sig.
groups	filled pauses	5	1.99	.088
	words per minute	5	3.59	.005*
	reformulation	5	2.84	.020*
	repetition	5	.42	.834
	replacement	5	1.10	.368
	false starts	5	1.52	.191

(*p<.05)

Post-hoc analyses reveal that the control group outperformed three post-task transcribing groups significantly in terms of speech rate, and at the same time it produced significantly more reformulations than two of the post-task groups. Effect sizes (Cohen's *d*) range from large to huge. The results are presented in Table 5.3 .

Table 5.3. Significant Results of Post-hoc Analyses on First Interactive Task: Fluency

dependent variables	groups	sig	effect size (*d*)
words per min	control > individual revision	.036*	1.40, very large
	control > pair revision	.004*	2.11, huge
	control > pair teacher revision	.051	1.0, large
reformulation	control > pair revision	.029*	1.06, large
	control > pair teacher revision	.029*	1.06, large

(*p<.05)

As we can see, in the first interactive task the control group produced more words per minute (i.e. speech rate) which means a more fluent speech. At the same time the control group used more reformulations which are the indicator for repair fluency (i.e. dysfluency). Since four repair fluency measures in total are considered in the present study, the adoption of more reformulations alone can not reflect the complete picture for dysfluency. In view of the significance concerning speech rate in favor of the control group, it is, therefore, assumed that the control group produced more fluent speech than a number of post-task groups in the first interactive task. Another one-way MANOVA was conducted for the second interactive task. The result showed no significant differences among the groups regarding fluency.

In exploring the effects of post-task transcribing on fluency, in the first interactive task, the control group produced a faster speech rate than certain post-task groups. It can be said that the baseline for further comparison is in favor of the control group. However, in the second interactive task there were not any significant differences between the control group and the post-task transcribing groups which indicated that the advantage of the control group which had emerged in the first interactive task did not exist any longer. This can be explained based on the descriptive statistics which show that there was an increase of speech rate among all the post-task transcribing groups from the first to the second interactive task performance. However, the control group produced fewer words per minute (i.e. a lower speech rate) in the second interactive task than it did in the first one. The control group, as a result, could not

maintain its advantage over the post-task groups. Without any significant difference noted in favor of the post-task groups in the second interactive task, we can only assume that the post-task transcribing groups have a tendency to produce their performance more quickly as a result of being involved in post-task transcribing, but no evident effects on fluency can be clearly suggested.

5.1.2 Complexity

We turn next to the more important form-linked components of performance. This section concerns the effects of post-task transcribing on complexity. The descriptive statistics show that in the first interactive task, the control group was at the middle level among all the groups as far as both complexity measures were concerned. However, in the second interactive task, the control group used the fewest clauses and words per AS-unit among the six groups. The results are given in Table 5.4.

Table 5.4. Descriptive Statistics on Interactive Task Performance: Complexity

Groups	Clauses per AS-unit				Words per AS-unit			
(n=16 for each)	DM1		DM2		DM1		DM2	
	mean	SD	mean	SD	mean	SD	mean	SD
individual	1.77	.14	1.88	.09	5.52	1.72	6.22	1.77
individual revise	1.70	.25	1.76	.15	6.07	1.61	6.87	2.02
pair	1.66	.13	1.90	.14	7.98	2.02	9.71	1.78
pair revise	1.63	.10	1.86	.15	8.16	1.12	9.47	2.03
pair teacher revise	1.74	.34	1.85	.13	6.91	.52	8.04	.60
control	1.67	.14	1.67	.14	5.70	.52	5.78	.60

One-way MANOVA results on the first interactive task performance show that there are significant differences among the groups: F(10,178)=4.48, p=.000, and the significant difference is reported for the number of words per AS-unit (p=.000), whereas there is no significant difference with respect to the number of clauses per AS-unit. The results are presented in Table 5.5.

Table 5.5. Univariate Tests on the First Interactive Task Performance: Complexity

source	dependent variables	mean square	F	sig.
groups	clauses per AS-unit	.042	1.024	.409
	words per AS-unit	21.048	8.842	.000*

(* p<.05)

Post-hoc analyses were conducted and the results revealed that the pair transcribing group (p=.001) and the pair-transcribing-and-revision group (p=.000) produced significantly more words per AS-unit than the control group in the first interactive task. Effect sizes are 1.6 and 2.9. The results are presented in Table 5.6.

Table 5.6. Significant Results of Post-hoc Analyses on the First Interactive Task: Complexity

dependent variables	groups	sig	effect size
words per AS-unit	pair> control	.001*	1.6, huge
	pair revision > control	.000*	2.9, huge

(*p<.05)

For the second interactive task performance, MANOVA results show significant differences among the groups: F(10,178)=9.827, p=.000 in terms of both the number of

clauses per AS-unit (p=.000) and the number of words per AS-unit (p= .000). Post-hoc analyses revealed that four (out of five) post-task groups employed significantly more clauses per AS-unit than the control group. In addition, three (out of five) post-task groups produced more words per AS-unit than the control group. The results are shown in Table 5.7.

Table 5.7. Significant Results of Post-hoc Analyses on the Second Interactive Task: Complexity

dependent variables	groups	sig	effect size (*d*)
clauses per AS-unit	individual > control	.000*	1.88, huge
	pair > control	.000*	1.72, huge
	pair revise > control	.001*	1.39, very large
	pair teacher revise > control	.002*	1.39, very large
words per AS-unit	pair > control	.000*	3.06, huge
	pair revise > control	.000*	3.56, huge
	pair teacher revise > control	.003*	1.56, huge

(*p<.05)

In sum, in the first interactive task, two pair-based transcribing groups produced more words per AS-unit than the control group. The baseline is said to be partially in favor of the post-task groups with respect to complexity. In the second interactive task, the superior performance of the post-task groups over the control group developed to a greater extent in that a) regarding words per AS-unit, in addition to the initial better performance of the two post-task groups, more post-task groups produced a larger number of words per AS-unit than the control group in the second interactive task with

larger effect size noted; b) more important, most (four out of five) post-task groups showed an advantage over the control group in adopting subordinations per AS-unit on which aspect all the groups were indistinguishable at the outset. Thus, the performance on the two measures of complexity is more or less consistent to suggest that the involvement of post-task transcribing has a significant impact on complexity in interactive tasks.

5.1.3 Accuracy

Accuracy—another crucial aspect to demonstrate the effects of focus on form activity—is examined in the present section. The descriptive statistics show that in the first interactive task, the control group was one of the groups which produced the fewest error-free clauses, shortest accurate clauses, and most errors per 100 words, while in the second interactive task, the control group yielded the fewest error-free clauses, shortest accurate clauses and most errors per 100 words as well. The results are show in Table 5.8.

Table 5.8. Descriptive Statistics on Interactive Task Performance: Accuracy

	error-free ratios				errors per 100 words				accuracy length			
	DM1		DM2		DM1		DM2		DM1		DM2	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
individual	.60	11	70	07	8 73	3 02	7 51	2 70	4 5	2 61	5 0	1 15
individual revise	63	12	78	05	7 37	2 70	6 03	1 91	4 0	1 63	5 63	1 78
pair	55	12	71	10	9.62	4 42	5 50	2 72	3 81	1 17	5 19	1 28
pair revise	56	08	77	05	9 17	3 47	6 21	2 72	3 44	1 09	5 06	2 05
pair teacher revise	58	16	76	.05	7 28	3 21	5 30	1 87	4 50	2 28	4 94	1 81
control	55	.06	56	06	10 30	2 86	9 84	3 14	3 44	1 15	4 56	1 71

However, one-way MANOVA results showed that in the first interactive task, all the groups were similar with regard to accuracy. In contrast, in the second interactive task, significant differences emerged between the control group and the post-task groups: $F(15, 243)=7.425$, $p=.000$ in terms of error-free ratio and errors per 100 words, whereas no significant differences are found concerning the accuracy length. The results are given in Table 5.9.

Table 5.9. Univariate Tests on the Second Interactive Task Performance: Accuracy

source	dependent variables	mean square	F	sig.
groups	percentage of error-free clauses	.114	26.63	.000*
	errors per 100 words	46.66	7.50	.000*
	accuracy length	1.925	.699	.626

(* $p<.05$)

Post-hoc analyses revealed that five post-task groups were significantly better than the control group with regard to error-free ratios. In addition, four (out of five) post-task groups produced significantly fewer errors per 100 words than the control group. The results are shown in Table 5.10.

Table 5.10. Significant Results of Post-hoc Analyses on the Second Interactive Task: Accuracy

dependent variables	groups	sig	effect size
ratio of error-free clauses	individual > control	.000*	2.22, huge
	individual revise > control	.000*	4.11, huge
	pair > control	.000*	1.88, huge
	pair revise > control	.000*	3.93, huge
	pair teacher revise > control	.000*	3.74, huge
errors per 100 words	individual revise > control	.001*	1.51, huge
	pair > control	.000*	1.61, huge
	pair revise > control	.001*	1.45, very large
	pair teacher revise > control	.000*	1.81, huge

(*p<.05)

To explore the effects of post-task transcribing regarding accuracy, the above results can be patterned together. At the outset, all the groups produced similarly accurate interactive task performance. We may say that the starting points for both post-task groups and the control group were generally the same. However, after the involvement of post-task transcribing, in the second interactive task, the similar performance among the groups was replaced by significant differences with respect to accuracy. In particular, *all* the post-task groups produced significantly larger ratios of error-free clauses with huge effect sizes noted. Besides, most post-task groups made fewer errors per 100 words than the control did. It is, therefore, evident that the involvement of post-task transcribing allows the post-task groups to gain significant

improvement in accuracy. Post-task transcribing, as a focus on form activity at post-stage, proves to be beneficial for language accuracy in interactive tasks.

5.1.4 Lexical Performance

When the formal aspects of task performance are considered, lexical performance is viewed as independent of complexity and accuracy (Skehan, 2009; Wolfe-Quintero et al, 1998). This section analyses the effects of post-task transcribing on the two lexical aspects of interactive task performance: lexical diversity and lexical sophistication. Descriptive statistics show that in the first interactive task, the control group, among all the groups, got the lowest scores in terms of lexical diversity and lexical sophistication; however, in the second interactive task, the control group got the highest score of lexical diversity but the lowest score of lexical sophistication among all the groups. The results are given in Table 5.11.

Table 5.11. Descriptive Statistics on Lexical Performance in Interactive Tasks

Groups	lexical diversity				lexical sophistication			
(n=16 for each)	N1		N2		N1		N2	
	mean	SD	mean	SD	mean	SD	mean	SD
individual	57.83	12.82	57.70	11.46	1.37	.25	1.64	.25
individual revise	54.66	12.36	54.39	11.53	1.49	.22	1.72	.14
pair	56.80	8.25	57.93	8.20	1.35	.21	1.54	.13
pair revise	53.35	10.96	57.46	9.97	1.36	.17	1.45	.15
pair teacher revise	53.84	7.44	56.58	9.55	1.33	.22	1.54	.21
control	50.75	9.28	60.96	11.13	1.24	.25	1.33	.23

One-way MANOVA results for the first interactive task show that there are no

significant differences among the groups which suggests that in the first interactive task, all the groups are fairly homogeneous in terms of lexical performance. In the second interactive task, the groups significantly differed on lexical aspects: $F(10,178)=4.01$, $p=.000$, and the significant difference is reported with respect to lexical sophistication only ($p=.000$) . The results are given in Table 5.12.

Table 5.12. Univariate Tests on the Lexical Performance in the Second Interactive Task

source	dependent variables	mean square	F	sig.
groups	lexical diversity	72.74	.675	.643
	lexical sophistication	.298	8.30	.000*

(* $p<.05$)

Post-hoc analyses revealed that four (out of five) post-task groups outperformed the control group in terms of lexical sophistication, that is the post-task groups produced significantly more infrequent words than the control group. Effect sizes range from large to huge. The results are presented in Table 5.13.

Table 5.13. Significances and Effect Sizes for Lexical Sophistication in the Second Interactive Task Performance

	treatment groups	sig	effect sizes & its level
control <	individual transcribing	.000*	1.33, very large
control <	individual transcribing and revising	.000*	2.12, huge
control <.	pair transcribing	.033*	1.16, very large
control <	pair transcribing and revising	.494	ns
control <	pair transcribing and teacher revising	.031*	0.98, large

(*$p<.05$)

In sum, in the first interactive task performance, the baseline for all the groups is equal. However, in the second interactive task, the involvement of post-task transcribing allows most post-task groups to be significantly superior to the control group by using more infrequent lexical elements, although it does not bring about such significant effects on lexical diversity. It may indicate that first, the two aspects of lexical performance are distinctive and independent of each other; second, post-task transcribing has significant effects on lexical sophistication, but not on lexical diversity.

5.1.5 Summary

The first research question attempts to explore the major concern in the present research: whether there are any effects of post-task transcribing on learners' task performance. In this chapter, the interactive task performance is the focus of analyses and the effects of post-task transcribing are explored at a general level by comparing all the post-task groups to the no post-task control group without consideration of the variety of post-task transcribing conditions (which will be the focus of the following sections).

As for the effects of post-task transcribing with regard to fluency, although in the first interactive task the control group produced more fluent speech than certain post-task groups, the advantage of the control group did not persist, that is to say, all the groups produced similarly fluent task performance in the second task. However, we should still be cautious to suggest any effects of post-task transcribing on fluency given that no significant differences were reported in favor of the post-task groups in the second interactive task.

When the effects of post-task transcribing are considered for complexity, the

analyses for the first interactive task performance reveal an advantage of two post-task groups over the control group in terms of mean length of AS-unit (i.e. words per AS-unit). As such, further comparison takes this preliminary advantage into consideration. In the second interactive task performance, the effects of post-task transcribing on complexity are clearly demonstrated in that most post-task groups produced significantly more subordinations per AS-unit than the control group, but they were indistinguishable in this performance area in the first task. In addition, as compared to the initial advantage in two post-task groups, superior performance for post-task groups emerged in terms of mean length of AS-unit. It may be said that the advantage of the post-task groups in the first task was extended.

More important, with regard to accuracy, the involvement of post-task transcribing leads to a striking effect on interactive task performance. Initially, all the groups were homogeneous in terms of accuracy in the first interactive task. But, later in the second interactive task, *all* the post-task groups demonstrated superior performance to that of the control group by producing a larger proportion of error-free clauses. In addition, most post-task transcribing groups produced significantly fewer errors per 100 words. It can be said that the post-task transcribing pushes learners to pay more attention to language form, in particular to produce more accurate speech in the interactive tasks.

Last, the effects of post-task transcribing on lexical performance are considered. While all the groups produced indistinguishable lexical performance in the first interactive task, most post-task transcribing groups proved to be superior to the control group in term of lexical sophistication in the second interactive task. Since the post-task groups did not generate better performance in terms of lexical diversity, it is suggested that post-task transcribing is beneficial with respect to lexical sophistication,

but not so on lexical diversity.

In sum, in the interactive task performance, the involvement of post-task transcribing activities leads to more complex and accurate language. In addition, post-task transcribing extends support to the employment of more infrequent lexical items. Given the encouraging achievement in language improvement in the present study, post-task transcribing proves to be an effective option to achieve the effects of focus on form at the post-task stage.

As mentioned above, the first research question is concerned about the general effects of post-task transcribing on learner's oral performance. In the following sections, research questions 2, 3, and 4 are addressed in an attempt to further explore the different effects of various post-task transcribing conditions in interactive tasks. The comparisons, therefore, are carried out among the post-task groups in the second interactive task.

5.2 Data Analysis regarding Research Question 2 and Research Question 3

In the present research, a variety of conditions are adopted in different post-task transcribing groups. Research question 2 concerns the relative effects of the individual- and pair-based post-task transcribing. Research question 3 is interested in the effects of further revision after post-task transcribing. As shown in Table 5.14, the two independent factors in research question 2 and 3 are interwoven among four post-task groups which constitute a balanced 2×2 design. To tease out an intervening factor—the teacher's role at post-task stage—the fifth group (i.e. the pair transcribing and teacher revision group) is excluded from the analyses in this section and will be considered in the next one.

Table 5.14. 2×2 design for research question 2 and 3

condition	individual	pair
no revision	individual transcribing group	pair transcribing group
revision	individual transcribing and revision group	pair transcribing and revision group

Two-way MANOVAs were performed to examine the effects of the individual/pair condition and the effects of revision as well as any interaction between them. The analyses are presented in terms of fluency, complexity, accuracy and lexical performance.

5.2.1 Fluency

This section concerns the effect of individual/pair-based post-task transcribing and the effects of further revision with regard to fluency. The descriptive statistics show that when the individual groups are compared to the pair groups, the pair groups used more filled pauses, repetitions and false starts than the individual ones, whereas the latter used more reformulations, but produced more words per minute than the former. When the revision and the no revision groups are compared, the no-revision condition produced more words per minute, and used more reformulations, repetitions and replacements than the revision condition, whereas the former used fewer filled pauses and false starts than the latter.

Table 5.15. Descriptive statistics of the four treatment groups in interactive tasks: Fluency

Conditions	filled pauses		words per min		reformulations		repetitions		replacements		false starts	
(n=16 for each group)	indivi dua-l	pair	indivi du-al	pair	individu al	pair	individ- ual	pair	individu -al	pair	individu al	pair
	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)
revision	2 63 (1.15)	3 57 (1 90)	67 99 (9 35)	71 98 (16 8)	94 (.68)	56 (.63)	4 31 (2 70)	5 44 (1 59)	1 19 (98)	81 (75)	31 (40)	25 (45)
no-revision	2.56 (2 45)	3 25 (2 18)	78.73 (7 58)	72 43 (15 6)	1 13 (1 15)	75 (93)	4 63 (3 34)	6 13 (1 54)	1 13 (1 02)	1 38 (1 31)	19 (40)	19 (40)

Two-way MANOVA results show that no interaction effect is reported between the two factors. In addition, there is no significant effect either of the individual/pair condition or of further revision. Thus, it is indicated that the differences between the individual/pair-based post-task transcribing and the presence of revision do not lead to any different interactive task performance in terms of fluency.

5.2.2. Complexity

In this section, the effect of the individual/pair condition and the effect of revision are examined with respect to complexity. The descriptive statistics (see Table 5.16) show that on the one hand, the pair groups produced more words per AS-unit and more clauses per AS-unit than the individual groups. On the other hand, the non-revision condition produced more clauses per AS-unit, but fewer words per AS-unit than the revision condition.

Table 5.16. Descriptive statistics of the four treatment groups in interactive tasks: Complexity

conditions	clauses per AS-unit		words per AS-unit	
	individual	pair	individual	pair
	mean (SD)	mean (SD)	mean (SD)	mean (SD)
revision	1.76 (.15)	1.86 (.15)	6.87 (2.03)	9.47 (1.40)
no revision	1.88 (.09)	1.90 (.14)	6.22 (1.77)	9.70 (1.78)

Two-way MANOVA reveals that first of all, there is a significant difference between the individual and the pair-based post-task groups: F (2,59)=23.68, p=.000. The univariate tests show that the pair condition produced significantly more words per AS-unit than the individual groups (p=.000, Cohen's *d*=1.76, a huge effect), whereas both individual and pair groups produced similar numbers of clauses per AS-unit.

In addition, a significance is reported concerning the effects of revision: F(2,59)=3.56, p=.035. The univariate tests show that the no-revision groups used more clauses per AS-unit than the revision groups (p=.014, Cohen's d=0.64, a medium effect), but the revision and no-revision groups did not significantly differ in terms of words per AS-unit. There is no interaction effect between the individual/pair condition and the involvement of revision.

Given that the significant effects of both factors are reported without any interaction between them, we will examine the results separately as follows.

As far as the effects of individual/pair-based post-task transcribing are concerned, the pair groups used more words per AS-unit than the individual groups, without any significant differences noted for clauses per AS-unit. This may suggest that the pair-based post-task transcribing has limited effects on complexity.

With regard to the effects of further revision after post-task transcribing, a *negative* effect of revision is reported among the groups. That is, the revision groups produced significantly fewer clauses per AS-unit than the no-revision groups. It can be said that in interactive tasks, the involvement of further revision, instead of promoting a more complex performance, may discourage learners from doing so, thus more simple language emerged in the interactive task performance. Unlike most of the previous research on revision which found it supportive in L2 language learning (Hyland and Hyland,2006; Ferris, 2003), this finding reveals a negative role of revision on complexity, and this will be further discussed in section 5.4 briefly and in Chapter 7 in detail.

5.2.3 Accuracy

This section concerns the effects of individual/pair work and the effects of revision on accuracy. The descriptive statistics show that for the comparison between the individual and pair based post-task transcribing, the pair condition produced fewer errors per 100 words, but shorter accurate clauses than the individual condition, while the two conditions got similar ratios of error free clauses. When the revision groups are

compared to the no-revision ones, the revision condition produced larger ratios of error-free clauses, lengthier accurate clauses, and fewer errors per 100 words than the no-revision condition. The results are given in Table 5.17.

Table 5.17. Descriptive statistics of the four treatment groups in interactive tasks: Accuracy

conditions	ratio of error-free clauses		errors per 100 words		accuracy length	
	individual	pair	individual	pair	individual	pair
	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)	mean (SD)
revision	.78 (.05)	.77 (.05)	6.03 (1.91)	6.21 (2.72)	5.63 (1.78)	5.06 (2.05)
no revision	.70 (.07)	.71 (.10)	7.52 (2.70)	5.50 (2.38)	5.00 (1.15)	5.19 (1.28)

A two-way MANOVA revealed that although there was no interaction effect between the two factors or no significant effects of the individual/pair-based post-task conditions, a significance was reported concerning the effects of revision: $F(3,58)=7.16$, $p=.000$, which was related to one of the accuracy measures: the proportion of error-free clauses. That is, the revision groups produced a significantly larger proportion of error-free clauses than the no-revision condition ($p=.000$, Cohen's $d=1.09$, a large effect). No significances are noted for the other two accuracy measures.

As we can see, the involvement of further revision after transcribing leads to a significant improvement in terms of accuracy, in particular to encourage the adoption of error-free clauses. It can, therefore, be indicated that to revise the task performance transcripts at the post-task stage brings about a clearly supportive impact on accuracy

in interactive tasks.

5.2.4 Lexical Performance

In this section, we look at the effects of individual/pair-based post-task transcribing and the effects of revision in terms of lexical performance. The descriptive statistics (see Table 5.18) show that with regard to the comparison between the individual and pair based post-task groups, the individual condition got a higher mean score for lexical sophistication, but a lower mean score for lexical diversity than the pair condition. As compared to the no-revision groups, the revision condition got lower mean scores for both lexical sophistication and lexical diversity.

Table 5.18. Descriptive statistics of the four treatment groups in interactive tasks: lexical performance

conditions	lexical diversity		lexical sophistication	
	individual	pair	individual	Pair
	mean (SD)	mean (SD)	mean (SD)	mean (SD)
revision	54.39 (11.53)	57.46 (9.97)	1.72 (.14)	1.45 (.15)
no revision	57.70 (11.45)	57.93 (8.20)	1.64 (.25)	1.54 (.13)

Two-way MANOVA results show that there is a significant difference between the individual- and the pair-based post-task transcribing in terms of lexical sophistication, but not on lexical diversity. In particular, the individual post-task groups employed significantly more infrequent words than the pair groups did (p=.000,

Cohen's d=1.08, a large effect). There is no significant effect of revision and no interaction between the two factors.

5.2.5 Summary

In this section, the independent factors in research question 2 and 3 have been considered at the same time in a series of two-may MANOVAs. Research question 2 attempts to look for the difference between the individual and the pair-based post-task transcribing in terms of the four aspects of interactive task performance. The analyses reveal partial effects of either condition with regard to syntactic complexity and lexical sophistication respectively. On the one hand, the pair condition encourages learners to adopt more words per AS-unit; on the other hand, the individual post-task transcribing favours the employment of more infrequent words. There were no significant differences between the two conditions on fluency or accuracy.

Research question 3 explores the effect of revision after post-task transcribing in the interactive task performance. The post-task revision has demonstrated its impact on the two important aspects of language form: accuracy and complexity, but in contrasting ways: a supportive role for accuracy, but a negative impact on complexity.

5.3 Data Analysis regarding Research Question 4

Research question 4 concerns the different effects of revision. The three revision groups are taken into consideration, and the major concern in this research question is

about the role of teacher in revision as compared to the role of students own involvement in revision. One-way MANOVAs were performed to examine the different aspects of performance in the interactive tasks.

5.3.1 Fluency

This section focuses on the effects of three revision conditions with regard to fluency. The descriptive statistics (see Table 5.19) show that among the three groups, the teacher revision group used most filled pauses and most false starts, whereas the pair revision group produced most words per minute and most repetitions, and the individual revision group used the most reformulations and repetitions.

Table 5.19. Descriptive Statistics for the Three Revision Groups regarding Fluency in the Interactive Task

conditions	filled pauses		words per min		reformulation		repetition		replacement		false starts	
(n=16 for each	N2		N2		N2		N2		N2		N2	
group)	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
individual revision	2.63	1.15	67.99	9.35	.94	.68	4.31	2.70	1.19	.98	.31	.48
pair revision	3.56	1.90	71.98	16.81	.56	.63	5.44	1.59	.81	.75	.25	.45
teacher revision	4.25	3.45	68.70	13.35	.50	.52	5.19	1.60	1.19	.83	.69	.48

One-way MANOVA results show that there are no significant differences among the three groups in terms of all the fluency measures which may suggest that the three different revision conditions have similar effects on fluency.

5.3.2 Complexity

This section concerns the difference of the three revision conditions in terms of complexity. The descriptive statistics show that the pair revision group produced most clauses and words per AS-unit, whereas the teacher revision group took the second place concerning both complexity measure. The results are given in Table 5.20.

Table 5.20. Descriptive Statistics for the Three Revision Groups regarding Complexity in the Interactive Task

conditions (n=16 for each group)	clauses per AS-unit		words per AS-unit	
	N2		N2	
	mean	SD	mean	SD
individual revision	1.76	.15	6.87	2.03
pair revision	1.86	.15	9.46	1.39
teacher revision	1.85	.13	8.04	2.03

One-way MANOVA results show that there is a significant difference among the groups on complexity: $F(4,88)=4.5$, $p=.002$. Further univariate tests show that a significant difference is reported with regard to the number of words per AS-unit ($p=.001$), but not on the number of clauses per AS-unit. Post-hoc analyses (Tukey HSD) reveal that the significant difference is between the individual revision and the pair revision groups ($p=.001$) which may be due to the different effects of individual or pair condition (for detailed analyses see section 5.2.2). When the teacher revision group is considered, the descriptive statistics show that this group was in the middle between two students groups, and MANOVA results show no significant difference between the

teacher revision group and both student revision groups.

It may be suggested that first, the teacher revision group did not show any advantage over the students counterparts with regard to complexity; secondly, the student-involved groups were significantly different from each other which reflected the differences between the individual and pair-based condition: the pair-based revision is more helpful for the adoption of words per AS-unit in the interactive tasks.

5.3.3 Accuracy

This section focuses on the relative effects of student revision and teacher revision on accuracy. The descriptive statistics show that the three groups got similar error-free scores. The two student-involved groups were similarly better than the teacher revision group in terms of the accuracy length. However, the teacher revision group produced the fewest errors per 100 words among the three groups. The results are presented in Table 5.21.

Table 5.21. Descriptive Statistics for the Three Revision Groups in the Interactive Task: Accuracy

condition (n=16 for each group)	error-free ratios		errors per 100 words		accuracy length	
	N2		N2		N2	
	mean	SD	Mean	SD	mean	SD
individual revision	.78	.05	6.03	1.91	5.63	1.78
pair revision	.77	.05	6.21	2.72	5.06	2.05
teacher revision	.76	.05	5.31	1.86	4.94	1.81

A one-way MANOVA was carried out and no significance was reported regarding

accuracy. The results suggest that the three revision conditions generate similar effects on accuracy.

5.3.4 Lexical Performance

Lexical performance is analysed in this section when the differences of three revision conditions are considered. The descriptive statistics show that the individual revision group got the highest mean score for lexical sophistication, but the lowest mean score for lexical diversity among the three groups, and the teacher revision group took second place for both lexical measures. The results are presented in Table 5.22.

Table 5.22. Descriptive Statistics for the Three Revision Groups regarding Interactive Task Lexical Performance

condition (n=16 for each group)	lexical sophistication		lexical diversity	
	DM2		DM2	
	mean	SD	mean	SD
individual revision	1.72	.14	54.39	11.53
pair revision	1.45	.15	57.46	9.97
teacher revision	1.54	.21	56.58	9.55

One-way MANOVA results show that there is a significant difference among the groups in terms of lexical sophistication, but no significance regarding lexical diversity. Post-hoc analyses revealed that the individual revision group was significantly better than the other two revision groups. The results are presented in Table 5.23.

Table 5.23. Significance among the Three Revision Groups: Lexical Sophistication in the Interactive Tasks

comparison groups	sig.	effect size
individual revision > pair revision	.000*	1.92, very large
individual revision > teacher revision	.012*	1.04, large

(*p<.05)

In terms of lexical sophistication, the individual revision seems to be superior to the other types of revision. We can say, in other words, that the involvement of teacher in revision did not have significant benefits on lexical sophistication. No significance is reported among the three groups in terms of lexical diversity.

5.3.5 Summary

Research question 4 explored differences among the three revision conditions, especially the role of the teacher revision. The results with different aspects of task performance consistently suggest that a) when the teacher revised the task performance transcripts for the students, this did not seem to have any significant benefits on the various aspects of task performance as compared to student-involved revision. Furthermore, with regard to lexical sophistication, the individual revision group outperformed the teacher revision group significantly; b) there are some differences between the two student-involved revision groups. When revision at the post-task stage

was conducted by the students, pair revision had a limited effect on complexity, and individual revision on lexical sophistication, which correspond to the findings concerning the effects of individual/pair-based condition at the post-task stage (c.f. section 5.2). This will be further discussed in Chapter 7.

5.4 Summary of post-task transcribing effects on both task performances

So far, we have analysed the effects of post-task transcribing on both the narrative (Chapter four) and the interactive task performance. This section summarizes the results and the findings on both task performances. The results are given in Table 5.24.

Table 5.24. Results for Research Question One to Four in Both Task Performances

		fluency			complexity		accuracy			lexical performance	
		FP	SR	RF	clauses /AS	words / AS	error-free ratio	err. /100 wds	acc.len	lex. sophistication	lex. diversity
RQ1: effect of post-task transcribing	narrative	ns	ns	ns	partial sig. (*)*	ns	sig. (*****)	ns	ns	partial sig. (*)	ns
	interactive	ns	ns	ns	sig. (****)	partial sig. (***)	sig. (*****)	sig. (****)	ns	sig. (****)	ns
RQ2: effect of individual / pair transcribing	narrative	ns	ns	ns	pair> individual	individual > pair	ns	ns	ns	interaction between individual & revision	ns
	interactive	ns	ns	ns	ns	pair> individual	ns	ns	ns	individual > pair	ns
RQ3: effect of revision	narrative	ns	ns	ns	ns	ns	ns	ns	ns	interaction with individual	ns
	interactive	ns	ns	ns	negative effect	ns	sig.	ns	ns	ns	ns
RQ4: effect of three revision conditions	narrative	ns	ns	ns	ns	ns	ns	ns	ns	individual > pair and teacher revision	ns
	interactive	ns	ns	ns	ns	ns	ns	ns	ns	individual >pair and teacher revision	ns

(Note: the number of asterisks indicates the number of superior treatment groups).

As far as the general effects of post-task transcribing are concerned, with regard to accuracy, the results in both the narrative and the interactive task performances pattern in the same way. In both task performance analyses, as compared to the non-post task group (the control group), the treatment groups show significant effects on accuracy. Furthermore, the results in the interactive task performance show a more consistent effect on accuracy: first, among the three accuracy measures, significance is reported with regard to one measure (the percentage of error free clauses) in the narrative tasks, however, in the interactive task performance, significances are reported concerning two accuracy measures (the percentage of error-free clauses and errors per 100 words); second, in the interactive task performance, in view of the significant results, larger effect sizes are noted as compared to those reported in the narrative task performance which shows that the magnitude of the treatment effect is larger in the interactive tasks than that in the narrative task performance.

When the post-task effect is examined with regard to complexity, the results from both the narrative and the interactive task performances show a supportive role of post-task activity, but to a different extent in different task types. In particular, in the narrative task performance, one treatment group outperformed the control group significantly with regard to one complexity measure (the number of clauses per AS-unit), while in the interactive task performance, most treatment groups are significantly better than the control group in terms of both

complexity measures. In line with the findings regarding accuracy, the results here reveal that the interactive task is more promising to demonstrate the effects of post-task transcribing on complexity.

With regard to lexical sophistication, the results are encouraging as well. In the narrative task performance one treatment group (the individual revision group) significantly outperformed the control group which may only suggest a partial effect of post-task transcribing on lexical sophistication. In the interactive tasks, most treatment groups are significantly better than the control group. It is evident that in the interactive task performance, the effects of post-task transcribing in terms of lexical sophistication are more significant than that in the narrative task performance.

No significances are found concerning fluency and lexical diversity in both tasks, which shows that there are no effects of post-task transcribing on those two aspects.

The above findings concern the general effects of post-task transcribing activities. The following focuses on the results regarding the effects of different conditions in post-task transcribing.

In the present study, post-task transcribing is performed either individually or in pairs. First of all, some consistent findings are revealed in both the narrative and the interactive task performances. The results show that the effects of the individual vs. the pair conditions are mainly related to language complexity which

can be further classified as syntactic complexity and lexical sophistication. In both tasks, the pair condition significantly promotes the use of more complicated syntax than the individual condition does. On the other hand, the individual condition consistently fosters the adoption of more sophisticated lexicon relative to the pair condition.

However, there are some inconsistent results in both task performance analyses. When fluency is considered, whereas no significant effect of either condition is found in the narrative tasks, the results from the interactive task performance analyses suggest a selective effect of individual transcribing by producing fewer repair fluency indicators in the interactive task performance. That is, individual transcribing tends to encourage a more fluent speech than pair transcribing does.

In both tasks, there is no significant effect of either the individual or the pair condition on accuracy and lexical diversity.

In addition to the division between the individual vs. the pair condition, the post-task transcribing in the present study was carried out with or without further revision. The analyses on both task performances show inconsistent even contrast findings regarding the effect of revision. In particular, in the interactive task performance, the involvement of revision results in a more accurate language, but a less complicated and more dysfluent performance than the non-revision condition. There is no significance reported with regard to lexical performance.

However, in the narrative task performance, except the interaction effect between individual transcribing and revision, no other significant effects are reported with regard to revision effects.

The role of the teacher's involvement in post-task revision is another focus of the present research. Whereas no significance is reported in analyses for both task performance in terms of syntactic aspects (i.e. accuracy, complexity and fluency), a significant result is noted in both tasks: lexical sophistication. The individual revision group outperformed the other two revision groups significantly in terms of lexical sophistication. This means that in the present research, the teacher involved revision shows no significant advantage over the learner involved revision, while on the other hand, a certain type of learner revision (i.e. the individual revision) is more effective than the teacher revision.

All in all, in the present study, post-task transcribing, as a focus on form activity at the post-task stage, proves to be effective in producing more accurate and complicated language in task performances. Furthermore, given that post-task transcribing is operated in various conditions, the distinctive role of the individual/pair transcribing and the effect of revision after transcribing are explored. Encouraging findings are reported, although they are not consistent or even in contrast in different types of task performances. Detailed discussion is presented in Chapter 7.

Chapter Six Data Analysis on the Effects of Task Practice and Task Types

6.0 Introduction

Building on the literature in post-task research in which participants have been generally involved in one or two task sessions (Skehan & Foster, 1997; Foster & Skehan, forthcoming; Lynch, 2001, 2007), in the present research participants are engaged over a longer period. Altogether five task sessions are involved—one training session and four experimental sessions. The adoption of multiple task sessions makes it possible to explore longer-term effects of both task practice and post-task transcribing. The present chapter focuses on the effects of multiple task practices. In addition, two different types of tasks are adopted in the present study which enables the exploration of task effects on learners' performance.

6.1 Data Analysis on Research Question 5 and Research Question 6

Research question 5 focuses on the effects of multiple task practices and research question 6 aims to explore the effects of different task types on various task sessions. Both the control group and the treatment groups are considered.. Since an additional factor—the effect of post-task transcribing is involved among the treatment groups, when the within-subject effects of task practice and of task types are explored, the treatment groups and the control group should be

examined separately. Given that all the groups are engaged in two task sessions for each task types, repeated measures MANOVAs are performed for the control group and the treatment groups respectively a) to compare the task performance in the first and the second sessions to explore the practice effect; b) to compare the different performances in different task types. The results are presented with respect to different aspects of task performance. For the within-subject analyses, Figures are used to demonstrate clearly the differences between the two sessions and the two task types.

6.1.1 Fluency

With regard to fluency, first, the control group is examined. Figure 6.1 shows that in both task sessions, participants in the control group produced more words per minute in the interactive tasks than they did in the narrative tasks. Figure 6.2-6.6 show that in both task sessions, the narrative task performances produced more filled pauses, reformulations, repetitions, replacements and false starts than the interactive task performances.

Figure 6.1. Mean Number of Words per Minute in the Control Group

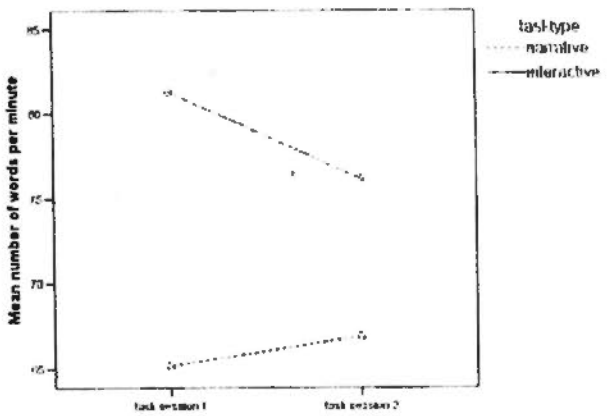


Figure 6.2. Mean Number of Filled Pauses in the Control Group

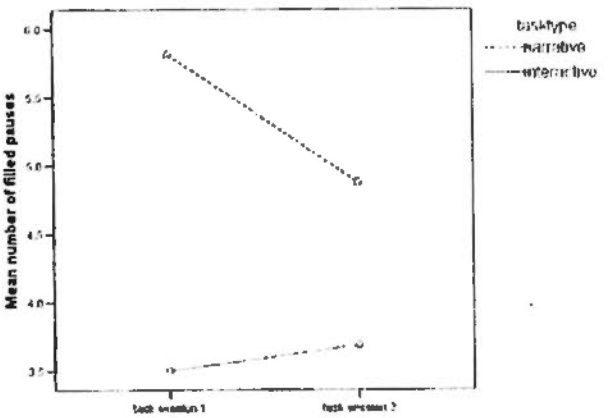


Figure 6.3. Mean Number of Reformulations in the Control Group

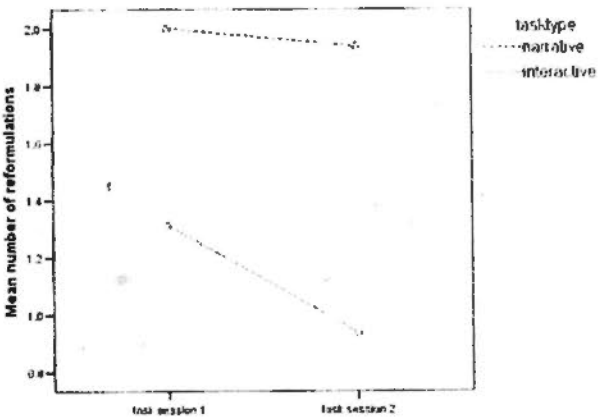


Figure 6.4. Mean Number of Repetitions in the Control Group

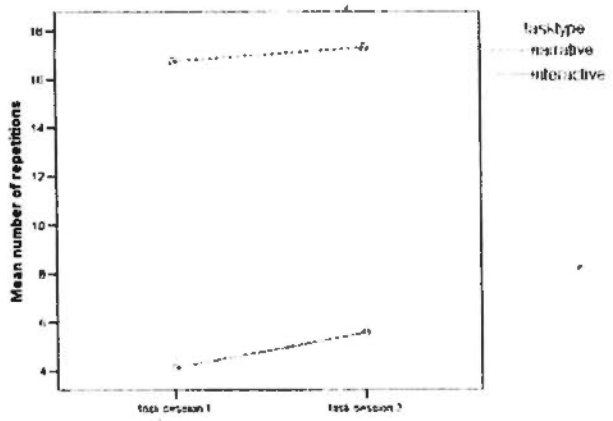


Figure 6.5. Mean Number of Replacements in the Control Group

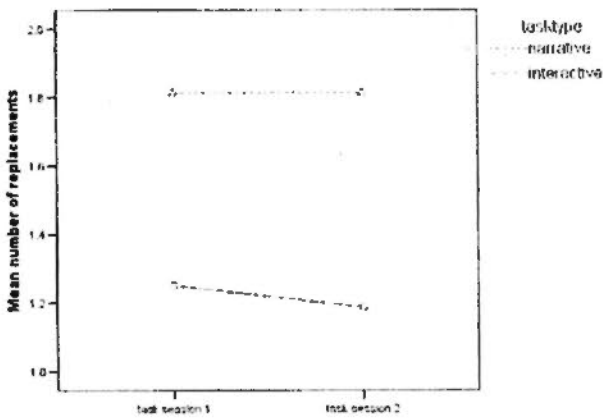
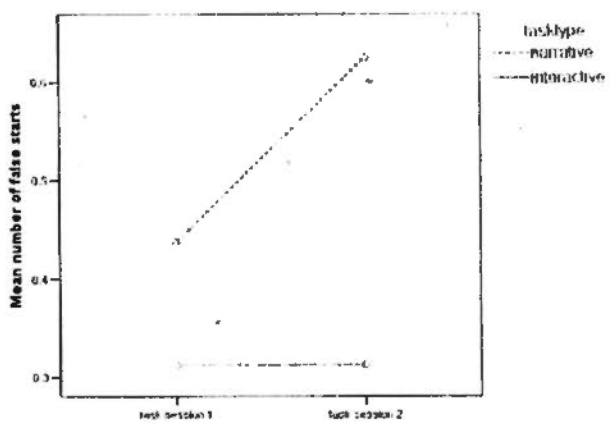


Figure 6.6. Mean Number of False Starts in the Control Group



The results of repeated measures MANOVA show that in the control group whereas there are no significant differences between different task sessions, there are significant differences with regard to task types: $F(6,10)=12.54$, $p=.000$. In addition, there is no significant interaction effect between task sessions and task types. Given the above significant result concerning the effects of task types, it is necessary to conduct further univariate tests to see which fluency measures are significantly different between the two task types. The results showed that with regard to words per minute, participants in the interactive tasks produced significantly more words per minute than they did in the narrative task performance; while with respect to a number of repair fluency and break down fluency measures, participants adopted significantly more filled pauses, reformulations and repetitions in the narrative task performance than they did in the interactive task performance. In view of the significant results, effect sizes (Cohen's *d*) were calculated, and the effect sizes were large or very large. The results are given in Table 6.1.

Table 6.1. Univariate Tests of the Differences between Two Task Types in the Control Group: Fluency

Source	dependent variables	F	sig.	effect size
task type	words per minute (I>N)**	14.35	.002*	1.38, very large
	filled pauses (N>I)	4.55	.050*	0.78, large
	Reformulations (N>I)	12.33	.003*	1.28, very large
	Repetitions (N>I)	27.89	.000*	1.93, very large

(* $p<.05$) (**: I refers to the interactive tasks; N refers to the narrative tasks)

In brief, the above results show that as far as the control group is concerned, the interactive tasks produced a higher speech rate (reflected by words per minute) and fewer repair and breakdown fluency indices than the narrative tasks. Both the higher speech rate and the fewer repair fluency indicators consistently reflect a more fluent speech in the interactive task performances.

Next, the treatment groups are considered. Figure 6.7 shows that there is a clear increase from the first session to the second session in terms of words per minute for both tasks. Furthermore in the first session, the narrative task performance produced fewer words per minute than the interactive task performance. However, in the second session, the former produced more words per minute than the latter. Figures 6.8-6.12 show that when the two task types are compared, in both task sessions participants in the narrative tasks produced more breakdown and repair fluency indicators than they did in the interactive tasks. When the two task sessions are considered, there are small increases in the second session in terms of reformulations, replacements and false starts in both tasks. There are slight decreases with respect to repetitions in the second session in both tasks. However, with regard to filled pauses, the narrative task performances used more filled pauses, whereas the interactive task performances produced fewer filled pauses in the second session.

Figure 6.7. Mean Number of Words per minute in the Treatment Groups

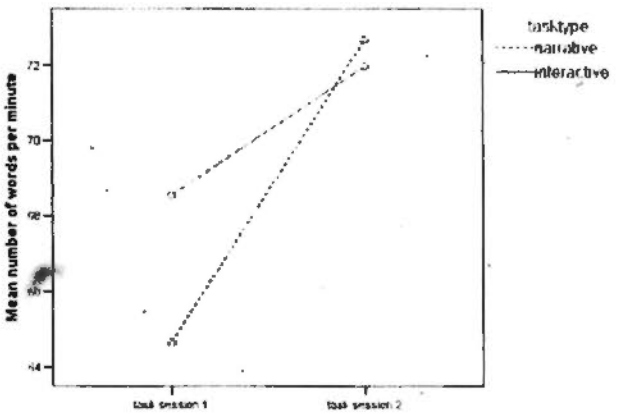


Figure 6.8. Mean Number of Filled Pauses in the Treatment Groups

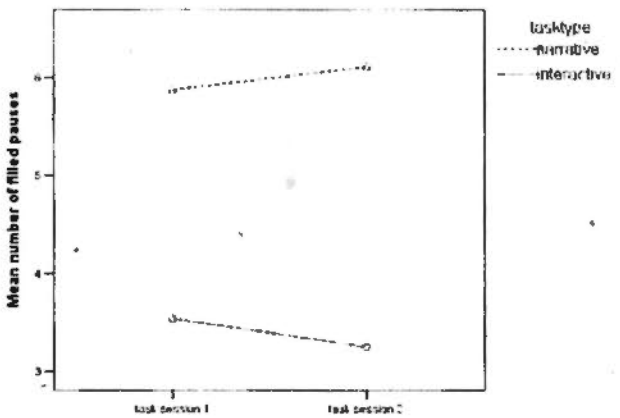


Figure 6.9. Mean Number of Reformulations in the Treatment Groups

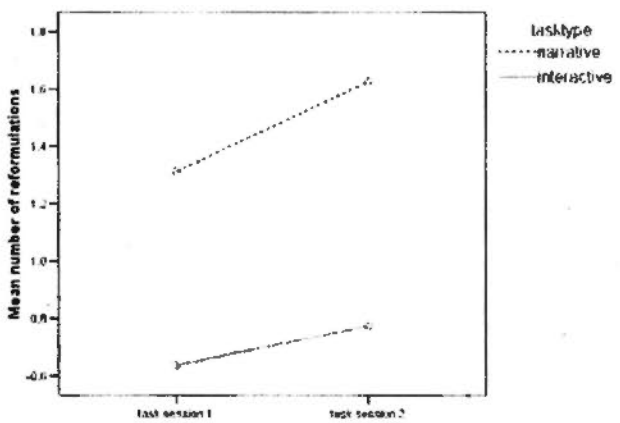


Figure 6.10. Mean Number of Repetitions in the Treatment Groups

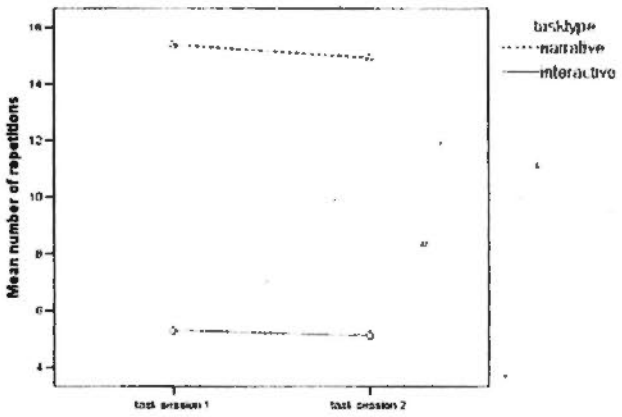


Figure 6.11. Mean Number of Replacements in the Treatment Groups

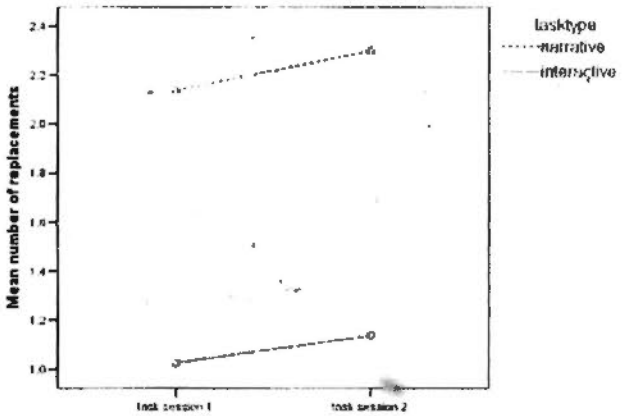
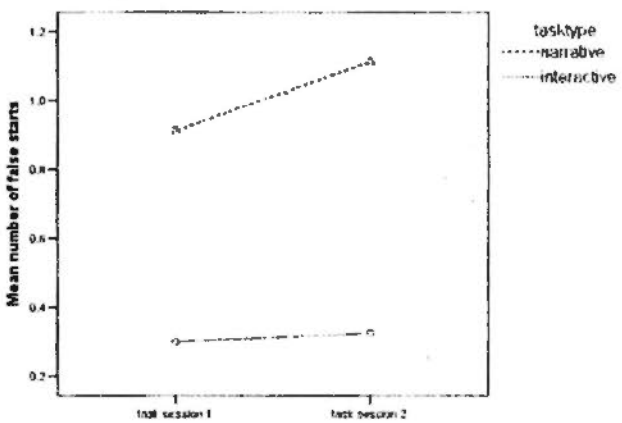


Figure 6.12. Mean Number of False Starts in the Treatment Groups



Repeated measures MANOVA shows that among the treatment groups there

are significant differences between the first and the second sessions: F(6,70)=4.27, p=.001. In addition, there are significant differences between the two task types: F (6,70)=29.50, p=.000. There are no significant interaction between task sessions and task types. Further univariate tests are carried out and the results show that for the effects of task practice, the second session for both task types produced significantly more words per minute (p=.000) and more reformulations (p=.046) than the first session. As far as the differences between task types are concerned, significances are reported with regard to all the fluency measures except one (words per minute). In other words, the narrative tasks produced significantly more breakdown and repair fluency measures than the interactive tasks which suggests that in the interactive tasks, the treatment groups produced a more fluent speech than they did in the narrative tasks. Effect sizes calculations are performed and Cohen's *d* values range from small to very large. The results are presented in Table 6.2.

Table 6.2. Univariate Tests of Within-subject Effects in the Treatment Groups: Fluency

source	dependent variables	F	sig.	effect sizes
Task session	words per minute	17.48	.000*	0.67, medium
	reformulations	3.98	.050*	0.32, small
Task type (N>I)**	filled pauses	32.65	.000*	0.91, large
	reformulations	31.92	.000*	0.90, large
	repetitions	103.81	.000*	1.62, very large
	replacement	54.37	.000*	1.17, very large
	false starts	40.26	.000*	1.01, very large

(* p<.05) (**: I refers to the interactive tasks; N refers to the narrative tasks)

Given the results regarding both the control group and the treatment groups, the following findings are reported: first, with regard to the effect of task types, the results consistently show that the interactive tasks promote a more fluent speech than the narrative tasks. Second, with regard to the effect of task practice, in the control group, no significant improvement is reported on fluency. However, when the treatment groups are considered, significant effects are reported for both task types to produce a more fluent speech in the second session than they did in the first session. In view of the fact that all the treatment groups are involved in multiple sessions of post-task transcribing and the control group shows no effect of simple task practice, it can be suggested that the improvement of fluency from the initial to the last session for both task types is due to the multiple post-task practices, rather than the mere task practice. One thing is worth mentioning here. Table 6.2 shows that with regard to the effects of task types, large or very large effect sizes are noted, whereas with regard to the difference of task sessions, the effect sizes are small or medium. It may further suggest that the effect of multiple post-task practices is not as strong as that of task types.

6.1.2 Complexity

This section focuses on the effects of task practice and of task types on complexity. The control group and the treatment groups are examined separately.

As far as the control group is concerned, Figure 6.13 and Figure 6.14 show

that when the two task sessions are compared, there are improvements in the narrative tasks in the second session in terms of both complexity measures, whereas there is no clear improvement in the interactive tasks in the second session with regard to both complexity measures. Furthermore, when the two task types are compared, Figure 6.13 shows that participants produced more clauses per AS-unit in the interactive tasks than they did in the narrative tasks, and Figure 6.14 reveals that in the latter participants produced more words per AS-unit than they did in the interactive tasks.

Figure 6.13. Mean Number of Clauses per AS-unit in the Control Group

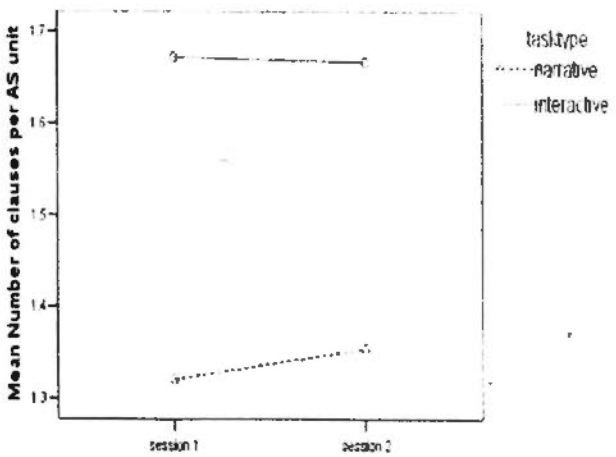
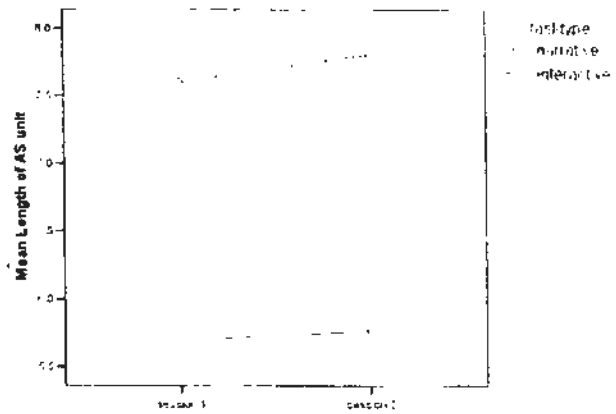


Figure 6.14. Mean Length of AS-unit in the Control Group



Repeated measures MANOVA results show that with regard to complexity, whereas there is no significant difference between task sessions, there is a significant difference between the two task types: F(2,14)=71.99, p=.000. In addition, no significance is reported for the interaction between task sessions and task types. In view of the significance with regard to the effects of task types, further univariate tests are performed and the results show that significances are reported with regard to both the number of clauses per AS-unit and the mean length of AS-units (words per AS-unit), with huge effect sizes noted. The results are given in Table 6.3.

Table 6.3. Univariate Tests of Within-subject Effects in the Control Group: Complexity

Source	dependent variables	F	sig.	effect size
task type	clauses per AS-unit (I>N)**	84.29	.000*	3.35, huge
	words per AS-unit (N>I)	33.33	.000*	2.11, huge

(* p<.05) (**: I refers to the interactive tasks; N refers to the narrative tasks)

In brief, the above results show that in the control group, different task types

have distinct effects on the two measures of complexity. While the interactive tasks pushes learners to produce more clauses per AS-unit, the narrative tasks encourages them to employ more words per AS-unit.

Then, the treatment groups are examined. Figure 6.15 and Figure 6.16 show that with regard to the differences between two task sessions, in both task types, participants in the second session got higher scores in both complexity measures than they did in the first session. With respect to the differences between two task types, the interactive tasks produced a larger number of clauses per AS-unit than the narrative tasks, whereas the latter produced more words per AS-unit (mean length of AS-unit) than the former.

Figure 6.15. Mean Number of Clauses per AS-unit in the Treatment Groups

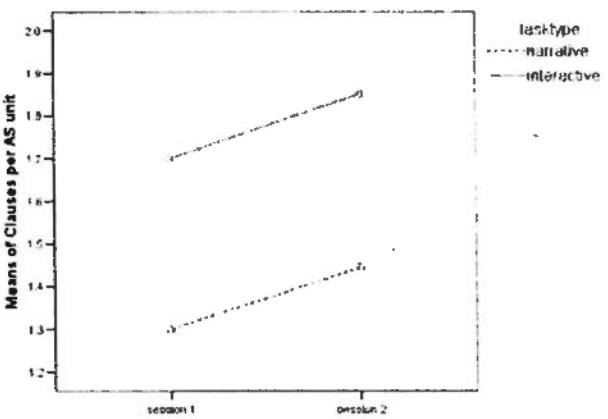
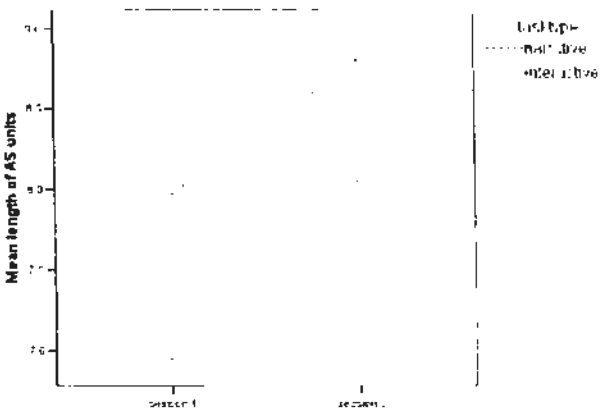


Figure 6.16. Mean Length of AS-unit in the Treatment Groups



A repeated measures MANOVA was performed and the results show that significances are reported with regard to both the effects of task sessions: F (2,78)=70.69, p=.000, and of task types: F (2,78)=203.89, p=.000. There is no significant interaction between task types and task sessions. Further univariate tests show that there are significant differences with regard to both complexity measures, with medium, very large or huge effect sizes noted. The results are given in Table 6.4.

Table 6.4. Univariate Tests of Within-subject Effects in the Treatment Groups: Complexity

source	dependent variables	F	sig.	effect sizes
task session	clauses per AS-unit	95.25	.000*	1.55, very large
	words per AS-unit	78.49	.000*	1.41, very large
task type	clauses per AS-unit (I>N) **	412.179	.000*	3.23, huge
	words per AS-unit (N>I)	13.89	.000*	0.59, medium

(* p<.05) (**: I refers to the interactive tasks; N refers to the narrative tasks)

In brief, when the treatment groups are considered, with regard to the effects

of task types, the interactive tasks show a significant advantage over the narrative tasks in promoting learners to produce more clauses per AS-unit, whereas the narrative task performance outperformed the interactive task performance in terms of mean length of AS-units. This finding is consistent with the results concerning the control group. In addition, with regard to the differences of task sessions, when the treatment groups are considered, there is a significant improvement in the second task session in both the narrative and the interactive task performances. The improvement in the treatment groups may be attributable to the effects of multiple post-task practices, rather than of mere task practices. In other words, due to the involvement of post-task transcribing on three occasions, learners are encouraged to use more complicated language in the both narrative and interactive task performance.

The results concerning both the control group and the treatment groups show that there are clearly significant but different effects of the two task types on complexity. In particular, while the interactive tasks encourage learners to use more clauses per AS-unit, the narrative tasks promote more words per AS-unit. However, there is no significant effect of mere task practice on complexity, since in the control group, the involvement of task practice is not beneficial to improve learners' language complexity. Given the complexity improvement in the second session among the treatment groups, it may be suggested that the involvement of multiple post-task practices is effective to promote learners' language complexity.

6.1.3 Accuracy

This section focuses on the effects of task practice and of task types on language accuracy. First, the control group is considered. Figures 6.17, 6.18, 6.19 show that when the two task sessions are compared, there is no clear improvement in the second task session as far as the ratio of error-free clauses is concerned. When the mean number of errors per 100 words is considered, there is a consistent reduction of errors in the second task session in both tasks. Furthermore, regarding the means of accuracy length, in both tasks, participants produced longer accurate clauses in the second session than they did in the first one. Both the reduction of the errors per 100 words and the increase to longer accurate clauses show that participants made progress in the second session with regard to accuracy. When the two task types are compared, with regard to the three accuracy measures, the Figures show that the interactive task performances are more accurate than the narrative task performances in that the former got a higher ratio of error-free clauses, fewer errors per 100 words and longer accurate clauses than the latter in both task sessions.

Figure 6.17. Means of Ratio of Error-free Clauses in the Control Group

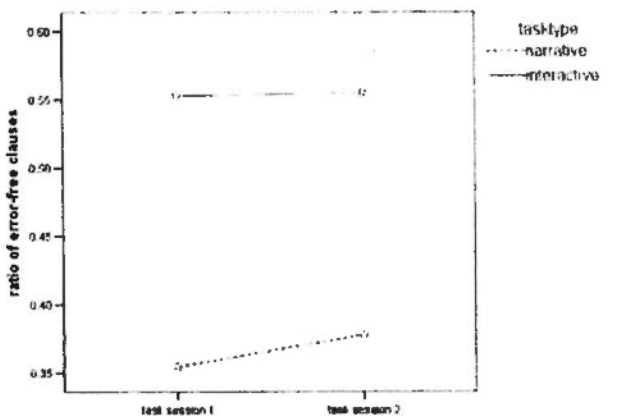


Figure 6.18. Means of Errors per 100 words in the Control Group

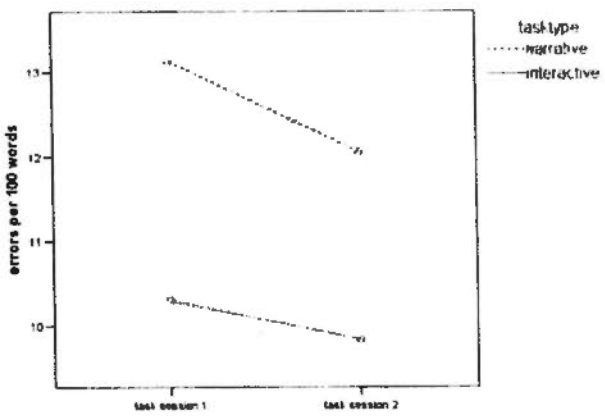
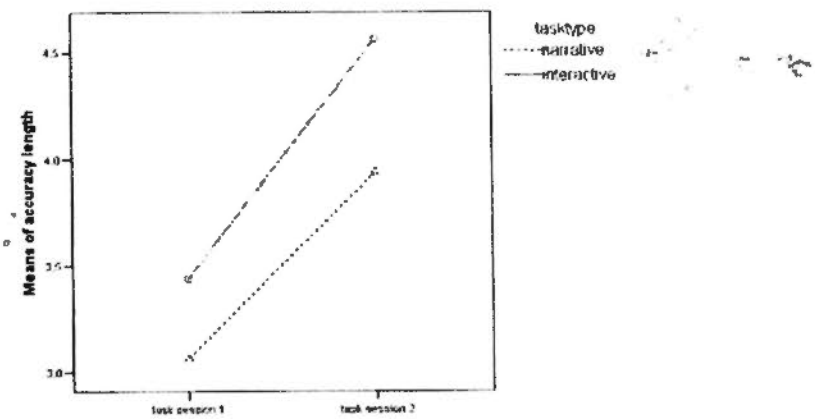


Figure 6.19. Means of Accuracy Length in the Control Group



A repeated measures MANOVA was carried out and the results show that with regard to the effect of task practice, for both task types, there is no significant difference between two task sessions. As far as the effects of different task types are concerned, significance is reported: F(3,13)=40.40, p=.000, whereas there is no significance concerning the interaction between task sessions and task types. Further univariate tests are performed to see which accuracy measures are significantly different. The results show that when the effect of task types is

concerned, significances are reported with respect to the ratios of error-free clauses and errors per 100 words, whereas no significance is reported concerning the accuracy length. Effect sizes were calculated and Cohen's *d* values are large or huge. The results are given in Table 6.5.

Table 6.5. Univariate Tests of Within-subject Effects in the Control Group: Accuracy

source	dependent variables	F	sig.	effect sizes
task type	ratio of error-free clauses (I>N)**	109.3	.000*	3.82, huge
	errors per 100 words (N>I)	4.79	.045*	0.8, large
	accuracy length (I>N)	.97	.341	ns

(* $p<.05$) (**: I refers to the interactive tasks; N refers to the narrative tasks)

The analyses on the control group show that whereas there is no significant effect of task practice on accuracy, there is a significant effect of task types. In particular, the interactive tasks encourage a more accurate use of language than the narrative tasks by producing more accurate clauses and fewer errors per 100 words.

Then, we take the treatment groups into consideration. Figures 6.20, 6.21 and 6.22 show that as compared to the first task session, participants in both tasks made progress by producing more accurate clauses, fewer errors per 100 words, and longer accurate clauses in the second task sessions.

When the comparison is between the two task types, Figure 6.20 and 6.21 show that with regard to the ratios of error-free clauses and errors per 100 words, the interactive task performances are more accurate relative to the narrative task

performance in that in the interactive task, participants used more error-free clauses and produced fewer errors per 100 words than they did in the narrative tasks. However, in Figure 6.22, with respect to accuracy length, the narrative tasks encouraged participants to improve more than the interactive task did, since in the first session, the narrative task performance was worse than the interactive task performance, but in the second session, the former was much better than the latter in terms of accuracy length.

Figure 6.20. Means of Ratio of Error-free Clauses in the Treatment Groups

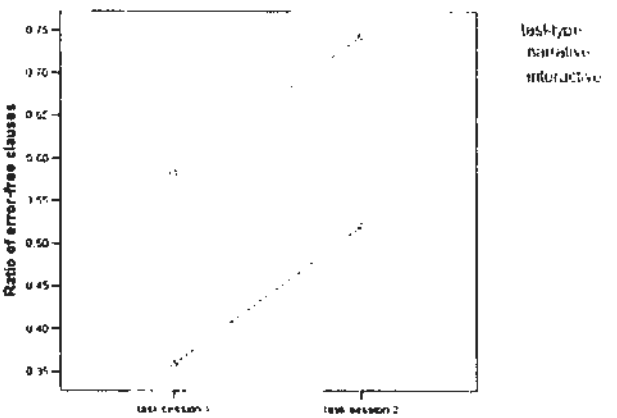


Figure 6.21. Means of Errors per 100 Words in the Treatment Groups

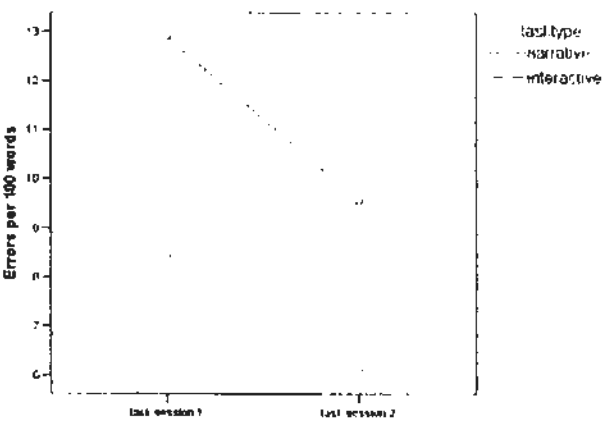
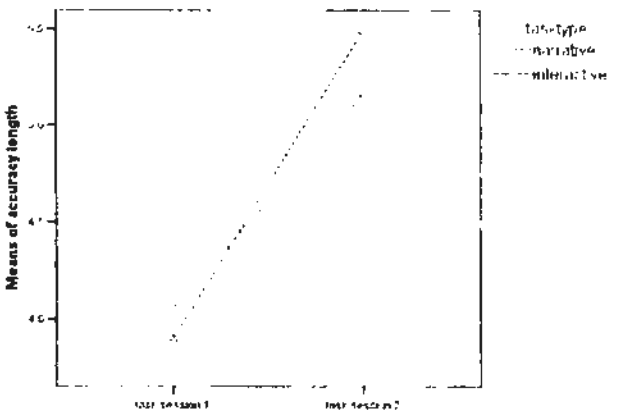


Figure 6.22. Means of Accuracy Length in the Treatment Groups



A repeated measures MANOVA was performed and the results show that there are significant effects of task practices: F=(3,77)=107.19, p=.000, and of task types: F=(3,77)=146.44, p=.000 on language accuracy in the treatment groups. However, there is no significant interaction between task practices and task types. Univariate tests were carried out and the results show that with regard to the effects of task practice, significances are reported concerning all the three accuracy measures, while when the effect of task types is considered, two accuracy measures (ratio of error-free clauses and errors per 100 words) show significances between two task types. The results are given in Table 6.6.

Table 6.6. Univariate Tests of Within-subject Effects in the Treatment Groups: Accuracy

source	dependent variables	F	sig.	effect sizes
task session	ratio of error-free clauses	324.25	.000*	2.87, huge
	errors per 100 words	68.46	.000*	1.32, very large
	accuracy length	34.05	.000*	0.93, large
task type	ratio of error-free clauses (I>N) **	318.40	.000*	2.84, huge
	errors per 100 words(N>I)	68.71	.000*	1.32, very large
	accuracy length (I>N)	.127	.723	ns

(* p<.05) (**: I refers to the interactive tasks; N refers to the narrative tasks)

The above results show that in the treatment groups, there are significant effects of task practice and of task types. With regard to the effects of task types, the treatment groups in the interactive tasks produced a more accurate language than they did in the narrative tasks. Furthermore, with regard to the effect of task practices, all the treatment groups, in both task types, improved language accuracy in the second task sessions.

In sum, the above analyses show that first, it is clear that there is a consistent effect of task types on accuracy in both the control group and the treatment groups. In particular, all the participants produced a more accurate performance in the interactive tasks than they did in the narrative tasks. Second, as far as the effect of task practices is concerned, the results from the control group and the treatment groups are not consistent. While there is no significant improvement in the control group as a result of the multiple task practices, there is significant enhancement in the treatment groups. Given that there is an additional factor in the treatment groups—the involvement of multiple *post-task* practices, we may suggest that the improvement of the treatment groups in terms of accuracy results from the effects of multiple post-task practice instead of mere task practice.

6.1.4 Lexical Performance

This section describes the analyses of the effects of task practice and of task types. Repeated measures MANOVA are performed for the control group and the

treatment groups separately.

When the control group is examined, Figure 6.23 and Figure 6.24 show that participants improved in the second session for both task types. When the two task types were compared, in the interactive tasks participants got higher scores for lexical diversity than they did in the narrative tasks, which meant that while doing the interactive tasks, participants did not repeat the used words as many as they did in the narrative tasks. On the other hand, in the narrative tasks participants got higher scores for lexical sophistication than they did in the interactive tasks which suggested that participants in the narrative tasks employed more infrequent words than they did in the interactive tasks.

Figure 6.23. Means of Lexical Diversity in the Control Group

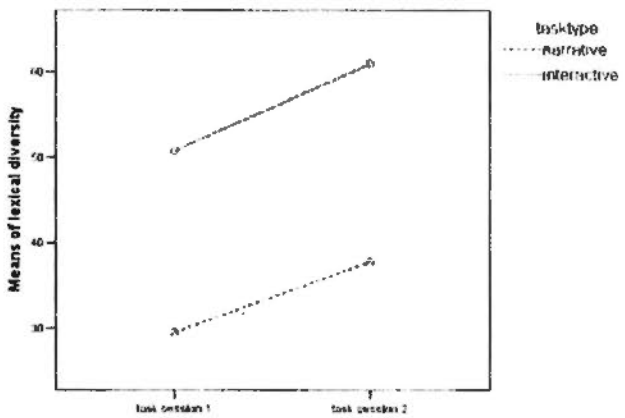
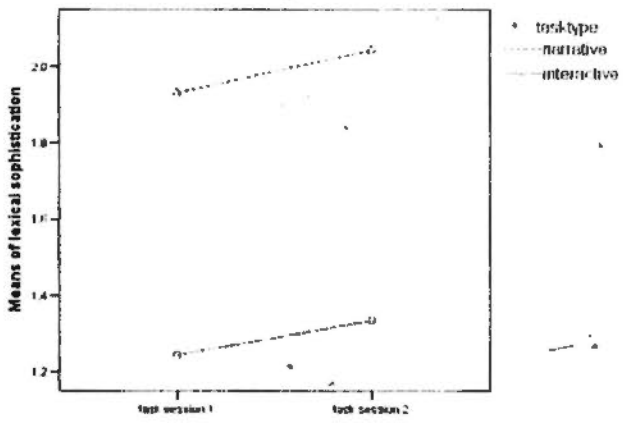


Figure 6.24. Means of Lexical Sophistication in the Control Group



A repeated measures MANOVA was performed and the results show that significances are reported for the effects of task practice: $F(2,14)=14.4$, $p=.000$, and of task types: $F(2,14)=46.14$, $p=.000$. There is no significant interaction between task types and task practices. Further univariate tests revealed that with regard to the effects of task practices and of task types, both aspects of lexical performance (i.e. lexical diversity and lexical sophistication) were significantly different. Effects sizes were calculated for the significant results and Cohen's d values were large or huge. The results are given in Table 6.7.

Table 6.7. Univariate Tests of Within-subject Effects in the Control Group: Lexical Performance

source	dependent variables	F	sig.	effect sizes
task session	lexical diversity	25.41	.000*	1.84, huge
	lexical sophistication	5.504	.033*	0.86, large
task type	lexical diversity (I>N) **	69.35	.000*	3.04, huge
	lexical sophistication (N>I)	35.01	.000*	2.16, huge

(* $p<.05$) (**: I refers to the interactive tasks; N refers to the narrative tasks)

The analyses show that there are significant effects both of task practice and of task type on lexical performance in the control group. The following are the analyses on the treatment groups. Figures 6.25 and 6.26 show that the two lexical aspects improved in the second session for both task types. When the two task types were compared, the interactive tasks got a higher score of lexical diversity than the narrative tasks, whereas the latter got a higher score of lexical sophistication than the former.

Figure 6.25. Means of Lexical Diversity in the Treatment Groups

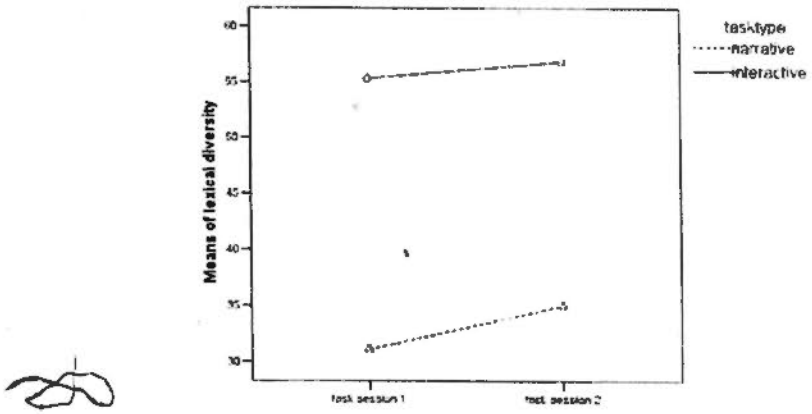
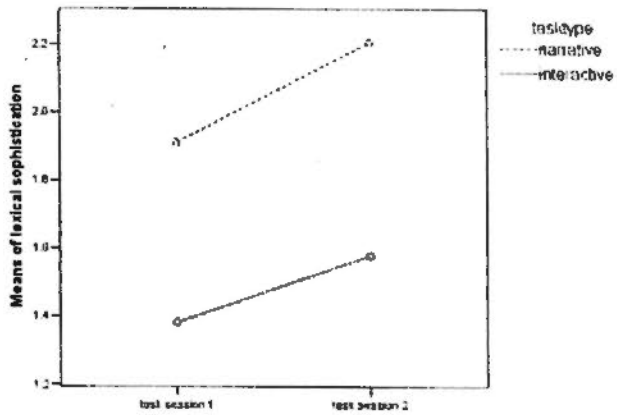


Figure 6.26. Means of Lexical Sophistication in the Treatment Groups



A repeated measures MANOVA was performed and the results show that there are significant effects both of task practice: F(2,78)=60.62, p=.000, and of task types: F(2,78)=478.84, p=.000). In addition, with regard to lexical sophistication, there is a significant effect of the interaction between task practice and task types: F (2, 78)=4.87, p=.010. Further univariate tests were conducted and the results show that both lexical diversity and lexical sophistication are significantly different when the effects of task practice and of task type are concerned. In view of the significant results, effect sizes were calculated and Cohen's d values ranged from small, medium to huge. The results are given in

Table 6.8.

Table 6.8. Univariate Tests of Within-subject Effects in the Treatment Groups: Lexical Performance

source	dependent variables	F	sig.	effect sizes
task session	lexical diversity	15.43	.000*	0.63, medium
	lexical sophistication	96.78	.000*	1.57, huge
task type	lexical diversity (I>N)**	551.43	.000*	3.74, huge
	lexical sophistication (N>I)	389.21	.000*	3.14, huge
task session * task type	lexical sophistication	5.86	.018*	0.39, small

(* p<.05) (**: I refers to the interactive tasks; N refers to the narrative tasks)

The analyses of the treatment groups show that when the effects of task practice are considered, the two task types have distinctive effects on both aspects of lexical performance. In particular, the interactive tasks promote significantly less use of repeated lexicon (or phrases) than the narrative tasks, while the narrative tasks encourage significantly more use of infrequent words than the former. Where the effects of task practice are concerned, participants improved significantly in the second task session in terms of both lexical aspects, although the effect of task practice on lexical sophistication is larger than that on lexical diversity, with larger effect sizes noted for the former. Furthermore, there is an interaction effect between task sessions and task types. This means that both effects are supportive to each other—the effect of task types is increased when more task sessions are involved.

In brief, the above results show that firstly, there are consistent effects of task practice on both the control group and the treatment groups with regard to both lexical aspects. As a result of being involved in multiple task practice sessions, participants significantly improved in terms of both lexical diversity and lexical sophistication. Secondly, there are clearly distinct effects of different task types on lexical diversity and lexical sophistication. In particular, the interactive tasks push learners to avoid repeating words to a larger extent than the narrative tasks do, while the latter contrive opportunities for learners to use more infrequent words than the former in task performance.

6.2 Summary of the Chapter

Research questions 5 and 6 concern the within-subject effects in the present study—the effects of task practice and of task types. The results are summarized in the following:

As far as the effects of task practice are concerned, the findings in terms of the syntactic aspects of task performance can pattern together. The control group was involved in multiple task practices, but no significant improvements were noted with regard to all the syntactic aspects of task performance. Given that significant improvements were reported for those aspects in the treatment groups which were involved in not only multiple task practice but also multiple post-task practice, it may be suggested that there are consistent effects of *multiple task and*

post-task practices, rather than of *mere multiple task practices* on fluency, complexity and accuracy.

In contrast to the above findings, the results on lexical performance show that there are effects of multiple task practices on lexical diversity and lexical sophistication. The two aspects of lexical performance improve in both the control group and the treatment groups. It can, therefore, be concluded that the involvement of multiple task practice encourages learners to use a variety of different words and to adopt more infrequent words in both interactive and the narrative tasks.

When the effects of task types are considered, the results with respect to fluency and accuracy consistently show a favorable effect of interactive tasks. In particular, interactive tasks promote a more fluent and accurate language than the narrative tasks do. However, the results regarding complexity and lexical performance reveal distinctive effects of different task types. In terms of complexity, the interactive tasks encourage learners to produce more clauses per AS-unit, while the narrative tasks push learners to adopt more words per AS-unit. This finding may be interpreted in another way: while in the interactive tasks, participants produce more clauses or clausal elements but shorter ones, in the narrative tasks, participants used longer but simple sentences.

In terms of lexical performance, the interactive tasks encourage learners to avoid repeating words and to use a variety of different words, whereas the

narrative tasks offer the opportunities for learners to employ more infrequent words in the task performance. This result shows that lexical diversity and lexical sophistication are the two distinctive aspects of lexical performance (Skehan, 2009) and different task types have distinct effects on those aspects. The above results are discussed in detail in Chapter 7.

Chapter Seven Discussion

7.0 Introduction

Based on the findings from the quantitative analyses on learners' task performance in the previous three chapters, this chapter attempts to interpret those findings from both theoretical and pedagogical perspectives in second language acquisition. The interpretation and discussion is presented according to the different research questions, in particular in terms of the effects of post-task transcribing, the effects of various conditions in transcribing, the influence of different task types, and the role of multiple task practice sessions on task performance. At the end, a summary table is provided to give a clear picture of the supportive experimental conditions for the different performance aspects, and the significant findings of the study are summarized as well.

7.1 Effects of post-task transcribing: get it right in the end

Most research on focus on form in task-based language teaching (TBLT) concentrates on its operation at the pre-task or during task stages. The present study chose a different perspective—the post-task stage—to explore the effects of focus on form in TBLT. In particular, at the post-task stage, participants were engaged in transcribing their own task performance. The primary concern of this study is to see whether the involvement of post-task transcribing is beneficial for the improvement of learners' task performance. The results reveal that

- *there are effects of post-task transcribing on formal aspects of task performance;*
- *the effects are not consistent in different aspects of task performance. In particular, the effects decrease in the sequence of accuracy, complexity, and lexical sophistication.*

7.1.1 General effects of post-task transcribing

To account for the effects of post-task transcribing, two issues may be related a) the foreknowledge of post-task transcribing; b) the operationalization of transcribing.

Prior to task performance, participants were informed that they would transcribe their performance recordings afterwards. It appeared that the foreknowledge of transcribing played a role in directing participants' attention to formal aspects of performance since the foreknowledge may remind them that "task performance is not an end in itself, but connects with wider pedagogic concerns" (Foster & Skehan, forthcoming). This connection emphasized the importance of the quality of task performance—to achieve a fluent communication, and at the same time to present a better performance. Participants, therefore, were cautious during performance to keep a balance between fluent communication and language forms. Accordingly, they may not only pay attention to meaning transmission for task accomplishment, but also allocate certain attention to the formal aspects of performance for a more satisfactory transcribed

performance, although it was not clear whether they shifted their attention unconsciously or intentionally.

The operationalization of post-task transcribing had effects on task performance as well. One of the most evident advantages of post-task transcribing is that it affords participants opportunities to attend to formal aspects of language performance. When doing tasks, most of the learners' attentional focus was inevitably on communication and meaning in order to get the task accomplished, while at the post-task stage, it was likely that more attention can be released to the formal aspects of the task performance, since meaning would not compete for major attention any longer (Foster & Skehan, forthcoming; Lynch, 2001, 2007). Noticing, as such, may occur more easily and naturally at the post-task stage which is a prerequisite for language change and acquisition (Schmidt, 2001).

However, much attention available for formal aspects only offers the *possibility* for noticing to take place, but does not guarantee its occurrence. In the present study the performance transcripts pushed participants to notice, remember, and reproduce the processed language forms (Lynch, 2007). On the one hand, transcripts transformed the oral task performance into written form which was beneficial for learners to attend to the performance since attentional capacity is increased when the information is presented in different modalities (Doughty & Williams, 1998a). On the other hand it reactivated the task performance even after its completion and thus offered a foundation for deeper processing, e.g. cognitive

comparison. "If the verbatim format of recent speech remains activated in memory and available for use in subsequent utterance formulation, this can be taken to be an important cognitive underpinning for facilitating the opportunity to make cognitive comparisons" (Doughty, 2001: 253). Cognitive comparisons may be made between the transcripts and the target language (i.e. noticing the gap) or between the missing forms in the transcripts and the existing counterparts in target language (i.e. noticing the hole) both of which functioned effectively for the improvement of formal aspects in the current research.

Further, post-task transcribing may be viewed as a type of task repetition. In a strict sense, it can be treated as a retrial of the task performance process, rather than a duplication of task performance. In terms of Levelt's Speaking model, task recycling of this sort may allow learners to exploit their familiarity on all the three components of speaking (Bygate, 2001), especially the conceptualization and the formulation stages. During task performance, participants were informed of the task type, task topic, task content and the corresponding task demands which were all concerned with the Conceptualizer. At the same time, they have already experienced the process of formulating the desired meanings. In transcribing, the need for meaning formulation has been minimized, thereby freeing attention for learners to attend to form (Gass et al, 1999; Lynch, 2007). This explanation is strengthened by the previous findings that task repetition promotes language performance in various aspects of task performance, such as accuracy, lexical

performance, complexity and language fluency (Bygate, 1996; 2001; Gass et al, 1999; Lynch & Maclean, 2000, 2001; Wang, 2009).

7.1.2 Different effects on various aspects of task performance

Post-task transcribing proved to be effective for language improvement. However, the focus on form effects are not consistent for different aspects of task performance, with the strongest effect of transcribing on accuracy, a more limited effect on complexity and lexical sophistication, and no effect on fluency and lexical diversity. What follows is a discussion of the different influences on those formal aspects.

In line with the previous post-task studies (Skehan & Foster, 1997; Foster & Skehan, forthcoming; Lynch, 2001, 2007), the present study reveals that accuracy is the best indicator to show the striking effect of post-task transcribing. Accuracy in language use can arise from three interacting sources: the degree of accuracy of the language representation itself; the strength of competing representations; and the degree of automatization of language production (Wolfe-Quintero et al, 1998). The first source of accuracy is dependent upon the established interlanguage (IL) system and long-term memory, and the other two pertain to the on-line real-time language use and rely on the allocation of attention. As cross-sectional rather than longitudinal research, the present study is mainly concerned with the factors related to on-line accurate language use, namely the competition between different

representations and the degree of automatization of language production. It is reasonable to assume that between the two competing representations (i.e. accuracy and complexity), participants focused most of their attention on the language conformity to the target forms and avoided errors in their attempts to achieve such a goal. Further, the multiple involvements of task practice and post-task transcribing increased the likelihood for language use to be automatized to some extent and consistently channeled learners' attention to the accuracy aspect of performance.

Despite the fact that most of participants' attention was directed to accuracy as a result of post-task transcribing, there are still certain effects on complexity and lexical performance, although the effects are not as strong and consistent as those on accuracy. Unlike the accuracy effect which occurred among all the post-task groups in both narrative and interactive tasks, the effects of post-task transcribing on complexity were shown in *all the pair-transcribing* groups and an individual-transcribing group in the interactive tasks and *one pair-transcribing* group in the narrative tasks. Thus, it may be inferred that it was the interactivity resulting from either the pair transcribing activity or the interactive task type, rather than the simple involvement of post-task transcribing itself, that generated positive effects on complexity. The role of interaction in the pair transcribing condition is discussed in section 7.2.1 and also in section 7.4.1 when the interactive and the narrative tasks are compared.

The current research reveals that focus on form at the post-task stage may be facilitative for the improvement of accuracy and complexity, although to different extents. In the literature, many of the focus on form studies are just concerned with the effects on accuracy since measures are generally in terms of target forms (Doughty & Williams, 1998a). In fact, the emergence of many intermediate IL forms that often represent increasing IL complexity rather than increasing accuracy (Skehan, 1996a), may be facilitated by focus on form as well. Doughty & Williams (1998c) argued that "Focus on form does not always immediately lead to IL changes that are reflected in increased accuracy. It may also lead to restructuring that reflects increased complexity, an equally important aspect of IL development. In this respect, focus on form has the advantage of affecting both IL development and IL accuracy" (p.254) .

At the same time, it is not surprising to see no effects of transcribing on fluency. Post-task transcribing is designed as a focus on form activity in the present study in attempt to explore its effects on the *formal* aspects of task performance. Fluency, which is more concerned with the *meaning* aspect of task performance (Skehan & Foster, 1997), was not expected to be influenced by the focus on form activity.

In the above, the discussion concerns the general effect of post-task transcribing, regardless of the differences between various transcribing conditions. In view of the pedagogical applications of this study, post-task transcribing was

operationalised through various conditions. The following sections present discussion on the different condition effects in greater detail.

7.2 Transcribing condition 1: pair-based vs. individual-based transcribing

The post-task groups fall into two categories: the individual-based and the pair-based transcribing groups. The effects of both of these conditions impact on one aspect of language: *complexity*, particularly syntactic complexity and lexical complexity (i.e. lexical sophistication). As far as lexical complexity is concerned, in both task types, the individual-based transcribing promoted a more complicated use of lexis, i.e. the use of more infrequent words. As far as syntactic complexity is concerned, results are not clear-cut concerning the two measures of complexity in both task types. In the narrative task, the pair condition produced more clauses per AS-unit than the individual condition; whereas in the interactive task the pair condition produced longer AS-units than the individual condition. Both results meant that the pair condition encouraged more complicated language than the individual condition did. However, an additional result blurred the above clear pattern: in the narrative tasks, individual transcribing, although resulting in fewer clauses per AS-unit, promoted longer AS-units than the pair condition. Possible explanations for the complicated results are provided in the next section, and then the role of individual transcribing on lexical complexity is discussed.

7.2.1 The role of interaction in pair-based transcribing

Despite the slightly contradictory result concerning the mean length of AS-unit in the narrative tasks which is discussed in the next section, in a most general sense, pair-based transcribing in both tasks encouraged a more complex syntax than the individual condition, although the effects on complexity were reflected in different measures.

From the perspective of cognitive SLA, restructuring may be the learning process which causes growth in complexity (Wolfe-Quintero et al, 1998). The process of restructuring "increases the chances that new forms will be incorporated into IL systems; promotes risk-taking and requires attention being devoted to the new forms of language which are being assembled" (Skehan, 1996:50). In this sense, unlike accuracy which is reflected by control over the internalized forms, complexity is more related to the internalization of new forms (Swain & Lapkin, 1995). There are at least two potential ways to promote IL restructuring: production of output plus metatalk, and production of output plus feedback (Swain, 1995; 2005). Both metatalk and feedback can result from interactions between speakers. It is also proposed by Skehan (1996a) that interactive opportunities are important to achieve IL restructuring and higher complexity.

From a psycholinguistic perspective, pair-based transcribing seemed to offer participants the opportunities to engage in the kind of moves which are facilitative

of L2 learning (Long, 1983, 1996), such as seeking and receiving confirmation, and providing each other with explanations about the original task performance. Swain & Lapkin (1998) point out that requests for confirmation about language form or language choice direct learners' attention to a specific language item. During pair transcribing, there were cases when participants could not fully understand what her/his partner exactly said in the recordings or why the partner said so. Requests for confirmation and the corresponding explanations may occur in the interaction. On the one hand the explanation-giver improved much from providing explanations concerning language forms because "it forces the learner to clarify and organize their own knowledge and thus enhances their own understanding" (Storch, 2007, p.155). On the other hand, the receiver of the explanation may also benefit since s/he was the person who first noticed the focused language forms, and also raised the check for confirmation and received the explanations. It was also possible that in pair transcribing, participants may give and receive immediate feedback on the earlier task performance, an advantage missing when participants transcribed individually. Joint responsibility over the creation of the transcripts means that students may be more receptive to peer suggestions and feedback comments (Storch, 2007).

7.2.2 Matching of transcribing conditions and task types

The above discussion had not accounted for the whole story concerning the

effects of pair transcribing on complexity. An additional result showed that whereas in the *interactive* tasks the *pair-based* transcribing groups produced longer AS-units than the individual condition, in the *narrative* tasks the *individual* transcribing condition promoted longer AS-units than the pair transcribing. It may be inferred that once there is a matching between task types and transcribing conditions (i.e. both performed individually or both in pair), there would be supportive effects on the mean length of AS-unit.

Dornyei's (2002, 2005) motivational task-processing system may be used to account for this matching effect. The motivational task-processing system consists of three interrelated mechanisms: task execution, task appraisal and action control. "These involve the students continuous monitoring and evaluating how well they are doing in a task, and making possible amendments if something seems to be going amiss" (Dornyei & Tseng, 2009, p.119). The transcribing activity functioned well for task appraisal.

In individual transcribing, participants may perceive the transcribing as self-reflection and task appraisal on his/her own earlier task performance. The negative signals from the task appraisal may trigger the need to activate action control processes, i.e. to apply the self-regulatory mechanisms for a better task execution in the subsequent tasks. In the narrative tasks, the action control was solely operated by the participant alone. In contrast, in the interactive tasks, due to the dynamic involvement of the partner, action control and task execution was not

determined by just one of the pair. Thus, the individual transcribing condition can only have effects on the narrative task performance, but not on the interactive one.

Things are different with the pair transcribing activity. When the pairs were reviewing and transcribing performance collaboratively, the transcribing may serve as a joint task appraisal not only of one's own performance, but also of one's peer's performance. The peer appraisals may threaten the "face" of the peer and cause a sense of peer competition, thereby bringing about different action control in different tasks. In the interactive task, the presence of a partner may strengthen the sense of "face saving" and of peer competition. As such the peers may pay more attention to action control, and contribute a better task execution. However, in the narrative tasks where the task was done with the absence of the partner, it may be speculated that the influence of the joint task appraisal was not as strong as that in the interactive tasks. As such, the individuals may not give as much attention to action control and task execution as they did in the interactive tasks. In this way, only the interactive task execution was promoted as a result of pair transcribing.

Given that the present study only focused on the quantitative aspects of task performance and no attention was given to the quality of transcripts and learners' reflection processes on task performance, the above discussion is mainly exploratory in nature. Further study would need to be conducted to examine the qualitative aspects of task appraisal and their relationship to task performance.

7.2.3 The role of intrapersonal communication in individual transcribing

Whereas pair-transcribing showed its effects on syntactic complexity, a fresh finding concerning the individual transcribing condition was noted. In both task types, individual transcribing promoted the use of more infrequent words, although this effect in the narrative tasks interacted with the role of further revision.

In the SLA literature, much research focus is on the effects of pair work (Lantolf, 2006; Long, 1983; 1996; Storch, 2001, 2005, 2007; Swain, 1995, 2005), and relatively little attention has been paid to the role of individual work. As for the comparison between the two conditions, some previous studies in first language acquisition found that pair and group work will not necessarily generate more gains relative to individual work. In one study, children working in pairs performed less well than individuals (Salomon & Globerson, 1989).

In the current research, the favorable effects of individual transcribing on lexical sophistication may be explained by the notions of *private speech* and *intrapersonal communication* within the sociocultural theoretical framework. When the individual participant transcribed the performance alone, without the presence of partners, s/he could only engage private speech for the clarification or modification of ideas, concepts and language focuses. Based on Vygotskian's theory of private and inner speech, intrapersonal communication is proposed by Vocate (1994) as a conversation whereby the I-you of social interaction becomes

an I-me dialogue, in which "I" makes choices on what to talk about and "me" interprets and critiques these choices (Lantolf, 2006, p. 96). In individual transcribing, when the narrative task performance was transcribed by the participant alone, essentially it was a process of self-reflection on the earlier performance. This self-reflection may be viewed as an "I-me" intrapersonal communication. "I" identified the idea or language focuses used in the performance and "me" critiqued these choices. When the interactive task performance was transcribed, the *interpersonal* communication during the tasks had an influence on the ongoing *intrapersonal* communication. The intrapersonal communication in the interactive performance transcribing, therefore, not only focused on the speech produced by the transcriber her/himself but also on the partner's speech, bringing about double benefits to the individual transcriber.

In contrast to the interactive task performance, in the narrative tasks individual transcribing alone cannot generate significant improvements in lexical sophistication, and it was the interaction between individual transcribing and further revision that brought about such effects on lexical sophistication. It seemed that it was more of a challenge to influence the narrative task performance by individual transcribing alone, since the participants were required to hold the floor in the whole process of task performance with little room left for online reflection and no opportunities to learn from their partner. However, when further revision is involved, the revision process can be viewed as a deep reflection on

earlier task performance and a reinforcement of the forgoing intrapersonal communication. The interaction between individual transcribing and revision has strengthened the effects of intrapersonal communication and provided participants with opportunities to use more infrequent words in the narrative tasks.

7.3 Transcribing condition II: further revision

7.3.1 With or without further revision after transcribing ?

In pedagogical practice, mere transcribing, although in different forms (e.g. individually or in pairs), may not be desirable because it inhibited learners' further reaction to the transcripts. Once learners are asked to transcribe their task performance, they tend to revise the transcripts to produce a better version by means of error correction, or text editing etc (Willis & Willis, 2007). The present study revealed that in the interactive tasks, the involvement of further revision after transcribing promoted learners' accuracy, but at the same time it had a negative effect on complexity.

These results are in alignment with the previous research in second language writing (Ferris, 2003; Hyland & Hyland, 2006) which shows that the involvement of revision has positive effects on language accuracy. When participants were asked to revise their transcripts, they were pushed to produce "modified output". As Swain (1993) put it, for interlanguage development, learners need to "reflect on their output and consider ways of modifying it to enhance comprehensibility,

appropriateness, and accuracy" (p.160-161). In the modifying and revising process, learners' primary concern was on the process of error correction which was related to a monitoring process for the control of forms (Ellis, 2008; Kormos, 1999). Based on Kormos' (1999) proposal for the role of monitoring in error-correction, the effects of revision on accuracy can be derived in the following ways.

i) Prior to correction, needing to judge whether it was an error or not (i.e. the identification of errors) channeled participants attention to reflect on the earlier performance, make use of their language resources and notice the gaps between their own IL system and the target language. In addition, the errors themselves, producing negative evidence in L2 language use, may push learners to alter their performance priorities by assigning greater importance to accuracy in task performance (Leeman, 2007), which may initiate the attempt to avoid the errors in task performances.

ii) Monitoring in revision made it more likely that learners' receptive knowledge would be converted into productive knowledge. The involvement of revision may redress the imbalance between the linguistic recognition knowledge and the linguistic production ability evident in their task performance. Therefore, it was reasonable to assume that monitoring during revision bridged the distance between

receptive and productive knowledge, and contextualised the newly-adopted productive knowledge as well. With the establishment of productive knowledge, accuracy was more likely to be enhanced in language use.

iii) When correcting errors, the participants rehearsed error-free solutions to language problems in short-term memory and this laid the foundation for the storage of the error-free solution in long-term memory for IL acquisition (Ellis, 2008). It may be inferred that the degree of the conformity of the L2 knowledge to the target language was enhanced by the inclusion of more error-free forms in long-term memory. Further, it was possible that participants might repeat the same errors in the task performance (e.g. – ed ending for simple past tense). The revision process made those errors perceptually salient, and the frequency and the salience of the errors may further promote the establishment of the correct IL system in long-term memory (Gass, 1997).

In contrast to the active role revision played on accuracy, a negative effect of revision was noted as well: once the participants had been involved in revision after transcribing, their language performance became less complex than that of the no-revision groups. This finding is in line with previous research which found that once learners paid attention to corrected errors, they tended to simplify their

writing to avoid situations in which they might make errors (Kepner, 1991; Sheppard, 1992). The existence of avoidance was identified in early error analysis studies which observed that learners who found a construction difficult tended to avoid it, using it only when especially confident that they can get it right, or when they had no choice (Truscott, 2007). This kind of strategy is termed the strategy of avoidance. It was possible that in the present study, due to the involvement of revision, the participants were more aware of the accuracy of the task performance. Responding to the error avoidance strategy, the counterproductive risk-avoidance strategy was used in the interactive tasks.

From a psycholinguistic perspective, the contrasting effects of revision on accuracy and complexity may be explained by the trade-off between these two performance areas (Skehan, 1998). In the present study the learners' primary concern during revision was on accuracy. Given the limited attentional resources available for the formal aspects of performance, complexity, as a competing area for attention, therefore received less attention during revision. In the task performance, participants tended to use simpler structures which they could control well without need for extra attentional resources. As such, more accurate but simpler performances were produced among the revision groups.

Given that the involvement of revision was not consistently supportive for different language aspects, whether or not to adopt revision in task-based language pedagogy may depend on the different goals of instruction. In contrast to

the revision effects on accuracy and complexity in the interactive tasks, further revision in the narrative tasks generated a favorable effect on lexical sophistication only when revision followed the individual transcribing. The interaction effect of revision and individual transcribing on lexical sophistication were discussed in detail in section 7.2.3.

7.3.2 Different revision conditions: individual, pair or teacher revision?

When the different conditions of revision (i.e. individual, pair or teacher revision) were compared, the present study noted that the only supportive role was provided by the individual revision condition on lexical sophistication. This result is in contrast to the previous revision research which showed favorable effects of pair revision or teacher revision on language accuracy in L2 writing (Ferris, 2004; Hyland and Hyland, 2006)

The account of the effect of *individual transcribing* on lexical sophistication (cf. section 7.2.3) may also be applied to explain the role of *individual revision* on the same lexical aspect since both of the activities were conducted *individually*. Put simply, it may be the case that the individual condition, either in revision or in transcribing, encouraged private speech and intrapersonal communication in the individuals. The individual revision group was engaged in the intrapersonal "I-me" communication which was basically focused on the lexical performance: to review and critique the use of words in the earlier task performance. Then,

relying on their own internal lexicon, participants retrieved more infrequent word as substitutes for the more general words.

Regarding the pair revision condition, it is expected that interaction during the pair revision process may be facilitative for the improvement of language performance. However, the participants in the present study were homogeneous in that they shared the same L1 background and were at the similar level of L2 proficiency. In general, it appears that "the less shared background interlocutors possess and the more heterogeneous the groupings are, the higher the amount of negotiation that can be expected in pair or group interactive practice" (Ortega, 2007, p.194; Varonis & Gass, 1985). It was likely therefore that relatively little negotiation and interaction occurred between the pairs in the revision process so that fewer opportunities for learning from peers arose, even though revision was performed in pairs.

To account for the insignificant effects of teacher-revision, some previous studies lend support to the result (Ferris, 2002; Lee, 2004; Lee & Schallert, 2008). Ferris (2002) found that students improved language accuracy as a result of either finding their own errors or making their own corrections, rather than receiving corrections from the teacher. What seems to be a crucial factor is having the students do something with the error correction besides simply receiving it (Lee, 2004). In the present study, when the transcripts were revised by the teacher and then were returned to the participants just for them to read the revised version,

there would not be any deep involvement of the revised transcripts which may only lead to surface processing of the revised transcripts. As a result, the teacher revision condition did not significantly influence task performance in the current study.

7.4 Task effects : interactive tasks vs. narrative tasks

In the current study, two different task types were adopted, and the comparison between performances in the two task types revealed that first, *interactive task performance produced more fluent and accurate language than the narrative task performance.* Second, *the differences between the two task types on complexity and lexical performance are mixed.* With complexity, the interactive tasks produced more clauses than the narrative tasks, whereas the latter produced longer AS-unit than the former. For lexical performance, the interactive tasks produced more varied lexis than the narrative tasks, whereas the latter used more infrequent words than the former.

7.4.1 The effects of interactive tasks on accuracy

Based on their series of task-based studies, Skehan & Foster (2001) proposed that interactive tasks are associated with greater accuracy and slightly greater complexity, but lower fluency. Results in the current research revealed similar results concerning accuracy and complexity, but different effects with fluency.

When language accuracy is concerned, interactive tasks encouraged more accurate language. In the dyadic communication, participants produced the L2 in the presence of their peers. To produce the second language is different from comprehending it in that whereas for L2 comprehension, semantic processing may be sufficient, for L2 production syntactic processing is unavoidable which is beneficial for language development (Swain, 1995). In interactive communication, participants were pushed to achieve control over syntactic and semantic processing for precise communication without much misunderstanding. Further, the speech produced by one partner served as the input for the other. The listener's input processing can alert him/her to potential gaps in his/her interlanguage system which required specific attention to language forms. As such, the IL system may be expanded by including more target norms or be improved by eradicating more erroneous forms, both of which were beneficial for the improvement of accuracy.

7.4.2 Task effects on different aspects of complexity

One point concerning the measure of complexity is noteworthy. In the present study, when complexity was measured in terms of clauses per AS-unit, the interactive task produced more clauses per AS-unit than the narrative task performance. However, when complexity is measured in terms of mean length of AS-unit, the results were in the contrast to the above—the narrative task

performance consisted of longer AS-units than the interactive performance.

The unexpected results, although in contrast to the assumption that longer AS-unit would consist of more clauses, can find support in other research on complexity. Byrnes, Maxim & Norris (in preparation), in a large-scale study concerning syntactic complexity development in L2 German writing, find that in both longitudinal and cross-sectional contexts, while the mean length of T-unit developed in a linear relationship with proficiency development, the number of clauses per T-unit improved at the first three levels in their study but *decreased* at the highest proficiency level. These results may suggest that the two measures of complexity do not always reflect the same aspects of complexity. Accordingly, Norris & Ortega (2008) propose that syntactic complexification occurs in distinct ways at different points in the development process, and multivariate measures of complexity are necessary corresponding to the multifaceted nature of the construct.

Regarding the present result, this may suggest that the number of clauses per AS-unit and the mean length of AS-unit do not necessarily positively correlate with one another. A longer AS-unit does not absolutely mean more clauses in that unit. In the literature, some researchers seem to hold the same position and have argued that length measures "do not discriminate between the various ways length can be achieved" (O'Donnell, 1976:33; Cooper, 1976; Yau, 1991; Wolfe-Quintero et al, 1998). Foster et al (2000) claimed that a lengthier turn can represent either

being better in productivity or in complexity where these two are different from one another.

So, it is worthwhile to explore the sources for the longer AS-units in the narrative tasks. In L2 writing research, researchers propose that there are two possible ways to account for the increase in T-unit length: "1) a writer can add more dependent clauses to the T-unit and 2) a writer can lengthen the existing clauses by adding phrases and words" (Cooper, 1976, p. 177; Wolfe-Quintero et al, 1998). Put it differently, words per T-unit can increase either through longer clauses with phrases and words or with more subordinate clauses. Kameen (1979, cited in Wolfe-Quintero et al, 1998) suggests that good writers produce longer T-units as a result of using more words rather than more clauses, most likely because they reduce clauses to prepositional, infinitive and participle phrases. In this same vein, in the current study, when participants were performing the narrative tasks (i.e. to describe the cartoon stories), they tended to use adjective, adverbial, and nominal phrases to refer to characters, motions and scenes. Therefore, they produced longer AS-units which included many phrases rather than clauses. On the other hand, when participants were involved in the interactive tasks, they attempted to reason and to deduce the solutions for the problems in which infinitives with clausal elements, adverbial clauses and relative clauses were more likely to be used. However, due to the limited available attention for on-line communication, they could only handle the shorter clauses or clausal

elements in the interactive tasks.

7.4.3 Task effects on different aspects of lexical performance

Similar to the distinctive task effects on syntactic complexity, the two task types had different effects on the two aspects of lexical performance. In particular, the interactive tasks produced higher lexical diversity, whereas the narrative tasks produced higher lexical sophistication, which echoes the findings from a meta-analysis of a series of task studies (Skehan, 2009).

In the narrative tasks, the only lexical resource for the participant is his/her internal lexicon, that is when describing a cartoon story, the participant who worked individually could only rely on his/her existing internal lexicon. However, the selection of words in the internal lexicon was constrained by the different cartoon episodes since the narrative task performance is input-driven and non-negotiable (Skehan, 2009). In other words, it is heavily reliant on the specific content of the narrative materials (e.g. pictures, or TV series). It may be inferred, therefore, that the sophisticated lexis in the narrative task performance was influenced by the two issues: the infrequent lexis which was embedded in the specific narrative task and the attempt to use advanced or infrequent words in the monologic context. As for the first issue, it was likely that certain advanced or less general words were implicated in the narrative tasks and were unavoidable to some extent when participants were describing the story in detail. As for the

second source, the attempt to use the less frequent words was related to the nature of the monologic tasks which were done individually. In section 7.2.3, the role of individual work and lexical sophistication was discussed in detail. In brief, the monologic tasks encouraged private speech, thereby resulting in intrapersonal communication concerning lexical use. This then pushed participants to adopt more advanced and less frequent words in the narrative tasks.

In the interactive tasks, more lexical sources are available for the participant. Due to the negotiable nature of interactive tasks (Skehan, 2009), in addition to one's own internal lexicon, the partner's speech afforded another source for lexical reference. As compared to retrieving from the advanced internal lexicon, noticing words from one's partner and using them online seemed to be more easily managed during communication. Another possibility was that partners' speech may have activated certain lexical items which already existed in the IL system. The participant did not need to learn from their partner, but relied on the triggered words in the interaction. Thus, the participant used a variety of words derived from his own mental lexicon or from the partner's speech.

Another feature of interactive tasks—unpredictability—also fostered the use of a variety of words in the interactive performance (Skehan, 2009). Different from narrative tasks adopted in this study in which the participants had the full knowledge of the story prior to the task performance (since participants watched the video before they retold the story), in interactive tasks there was no fixed

direction for the dyadic communication to develop, i.e. the upcoming task performance was unpredictable. A flexible communication flow may then make participants less likely to repeat the same sets of words in the interactive tasks so that a higher lexical diversity occurred in them.

The two lexical measures (i.e. lexical sophistication and lexical diversity) have been found to be independent of one another. Previous research (Laufer, 1994; Skehan, 2009) has also revealed no correlation between a measure of lexical sophistication and a measure of lexical diversity. It is inferred, therefore, that the two measures measure different aspects of lexical performance. As Laufer (1994) claimed, lexical variation (i.e. lexical diversity) does not measure how large the lexicon is, only how varied it is for that particular individual, but lexical sophistication does measure the size of the lexicon against an external standard. Different features of the two task types therefore had different influences on the two aspects of lexical performance.

7.5 Task practice effects

The participants in the present study were engaged in four task sessions. Although it may not be treated as a longitudinal study, the present study is different from the others in this area in that it attempted to explore the effects of being involved in multiple task sessions, i.e. a practice effect in task-based language teaching. The results showed that i) although the control group did not

receive any post-task treatments, the multiple task sessions promoted lexical performance among both the control group and the post-task groups; ii) Whereas there were no significant effects of mere task practice on other performance aspects in the control group, there are effects of *multiple post-task practice* on other aspects among the post-task experimental groups.

Levelt's Speaking Model (1989) may account for the practice effect on lexical performance. The multiple involvements of task practice gave participants ample opportunities for conceptualizing, formulating, and articulating in contexts where they were encouraged to convey messages relying on their existing lexical, grammatical, and phonological knowledge. In Levelt's Speaking Model, lexis takes precedence over grammar during formulation in that lexical items are the primary generator in promoting grammatical encoding in the L2. Thus, it may be the case that participants prioritized lexical performance during task practice.

The priority towards lexical items benefited the two lexical aspects in different ways: As far as lexical sophistication is concerned, multiple task practice familiarized the participants to some extent with the conceptualization of the messages. This may leave more attentional resources for lexical selection in the formulation process. As such, when participants were searching for the lexis at the formulation stage, they tended to access more infrequent or advanced words for grammar encoding and syntactic building. As for lexical diversity, the multiple task practices afforded participants opportunities to perform at the formulation

stage with less control (i.e. to become more automatic) so as to free more attention for learners to make decision for online lexical use by avoiding recycling of words in the tasks.

In sum, it may be inferred that although multiple task sessions generated significant effects on both aspects of lexical performance, different underlying speaking components were involved for the different aspects.

Except for the effect on lexical performance, there was no task practice effect on other performance aspects. Levelt (1989) claims that a lexical items' lemma contains not only semantic information but also syntactic information. A lemma is activated when its meaning matches part of the preverbal message, and certain syntactic building procedures would be activated at the same time. Levelt regards this syntactic building process as procedural knowledge for L1 speakers. But for L2 speakers, it is evident that the syntactic building processes will not become proceduralized without sufficient practice. In the present study, participants were engaged in four sessions which were far from sufficient for the proceduralization of the syntactic building processes. Given that both accuracy and complexity are related to the syntactic aspects of language processing, mere task practice, of a relatively short duration, could only lead to improvement at the lexical level. In particular this concerns the use of more infrequent words (which were mainly concerned with the semantic information in the lemma) or less recycling of sets of words.

However, among the post-task groups, multiple post-task practice *plus* multiple task practice did lead to improvement of some syntactic aspects in that the post-task practice, unlike the meaning-oriented task practice, channeled participants attention to form in an intentional way. Multiple post-task practice, therefore, afforded participants opportunities to gradually proceduralize the matching between the preverbal message and the syntactic building process, and this promoted a more automatic formulation stage in speaking. Therefore, the syntactic-based aspects of performance (i.e. accuracy and complexity) were improved as a result of multiple post-task practices.

7.6 Summary and synthesis of the findings

The present research establishes that the post-task stage has provided desirable opportunities to achieve a focus on form effect. Specifically, post-task transcribing directs learners' attention to the formal aspects of language. In addition, this study explored the benefits of multiple task practices, and also the different task effects of narrative and interactive tasks. The multifaceted design of the study brought about interesting results. Table 7.1 summarizes the supportive experimental conditions for the different aspects of performance.

Table 7.1 Supportive experimental conditions for different aspects of task performance

	Supportive Experimental Conditions			
Task Performance	post-task transcribing	transcribing condition	task effect	practice effect
accuracy	Yes	revision	interactive	No
complexity				
clauses per AS-unit	Yes	pair trans. in narrative	interactive	No
mean length of AS-unit	Yes	not consistent	narrative	No
lexical sophistication	Yes	individual transcribing & revision	narrative	Yes
lexical diversity	No	No	interactive	Yes
fluency	No	No	interactive	No

This chapter accounts for the results based on different theories, including both cognitive psychology and sociocultural theory. From the stand point of learning theory, the results of the present study generally indicated a series of significant findings, some of which are in line with the previous research, while the others may be regarded as new in the field:

Findings in alignment with the literature:

- The adoption of post-task transcribing is efficient for different formal aspects of task performance. The foreknowledge and the operationalization of post-task transcribing are responsible for the

benefits demonstrated in that they provide opportunities for noticing and deep processing in participants' language performance.

- Interactive tasks proved to be more promising for better overall task performance. The interaction engaged in by the pairs promoted collective scaffolding for each other and provided opportunities for negotiation of meaning, as well as negative feedback for language improvement.

New findings from the current research:

- The individual condition at the post-task stage led to an improvement in lexical sophistication, no matter whether the individual condition was concerned with transcribing, revision or task types (i.e. the narrative task). This effect may be due to private speech and the intrapersonal communication resulting from the individual work.
- Revision has complex effects on accuracy and complexity. The monitoring process involved in revision encouraged a positive effect on accuracy, whereas a risk avoidance strategy may account for a negative effect on complexity.
- The only effect of task practice is on lexical performance. Concepts in Levelt's Speaking Model which are concerned with the role of lexical encoding and syntactic encoding provide possible explanations for such a result.
- Measures of task performance need to be multivariate, since the construct underlying certain performance aspect, e.g. complexity and lexical performance, are multi-faceted. Different measures may reflect distinct features of the same construct.

In sum, it is fair to say that this research, on the one hand provides convincing evidence for the adoption of post-task activities (e.g. transcribing) in task-based instruction to achieve a focus on form effect. On the other hand this study affords new insights concerning task-based research and pedagogy which will be presented in the next conclusion chapter.

Chapter Eight Conclusion

8.0 Introduction

The effect of post-task activities has been unfortunately neglected by most researchers. The supportive role of post-task transcribing demonstrated in the present research provides striking evidence to show its importance, especially for a focus-on-form effect. The findings are particularly interesting in relation to task-based language teaching which emphasizes meaning transmission as a primary concern, as well as focus-on-form research which pays most attention to pre- and during-task activities.

In this conclusion chapter, first of all, the theoretical and the practical significance of the present research are explored. Then the pedagogical implications for task-based classrooms are considered. Based on the analysis of the limitations of the study, several suggestions are proposed for further studies concerning post-task focus on form research.

8.1 Significance of the Research

This task-based research study, which has examined post-task focus-on-form activity, is of both theoretical and practical significance. Transcribing, which can be implemented in a variety of ways, can shed light on the contribution that post-task activities can make to a focus on form effect. The consistent effect on accuracy among most post-task groups proves that post-task transcribing is

facilitative for second language improvement. Further, the different conditions of transcribing led to significant progress in various formal aspects. In accordance with previous related studies (Skehan & Foster, 1997; Foster & Skehan, forthcoming; Lynch, 2001; 2007), the present study argues that the inclusion of a post-task stage is indispensable in a task-based framework and studies on post-task activities are of valuable significance in task-based focus on form research.

This study affords insights about the role of noticing in a broader sense. In the literature on SLA, the role of noticing has been researched in terms of either input processing (Schmidt & Frota, 1986) or output production (Swain, 1995; Swain & Lapkin, 2001) separately. In the present study, during post-task transcribing, noticing input and output took place simultaneously. The task performance recordings served as the input, while the production of transcripts acted as the output (output-as-process and output-as-product) for noticing and further processing. The incorporation of both input and output perspectives brought the research focus to be closer to the overall process of L2 use.

The study provides data that can be used to address such interesting issues as the nature of interaction in pair work, the effect of private speech and intrapersonal communication in individual work, and the underlying process of revision. The role of interaction has received a great amount of attention in both cognitive psychology and socio-cultural theory with no consensus on its effect.

The present study specified the effects of interaction on syntactic complexity. At the same time, it revealed a new finding concerning the effect of individual work on lexical sophistication. Both individual work and the lexical aspects of task performance are relatively little researched in the task-based literature. This finding may stimulate further examination on those aspects. In addition, the involvement of further revision after transcribing brought about contrasting effects on accuracy and complexity, resulting from the monitoring process and counterproductive communicative strategy use.

This study takes into account the multifaceted nature of performance constructs, especially of complexity and lexical performance. In previous task-based research which measures performance in terms of accuracy, complexity, fluency and lexical performance, only a few studies have taken multivariate measures for one construct so as to explore its multifaceted nature. This study revealed, although as a *post hoc* finding, that the two measures of complexity (i.e. clauses per AS-unit and words per AS-unit) do not necessarily correlate with each other positively. More interestingly, it was found that different task types may impact different complexity and lexical measures in a contrasting way.

The research is of obvious practical relevance to language teachers, as it sheds light on how teachers can incorporate post-task transcribing in a task-based pedagogy, and how they can manipulate the conditions of transcribing for different performance achievement. Specifically, the research provides important

clues about the procedures teachers need to adopt with post-task transcribing in task-based classrooms. It also affords insights into how learners can achieve a balance between the different formal aspects of task performance in different post-task conditions. Moreover, the procedures and findings of this research can be generalized to regular L2 task-based classrooms, since a sufficiently large number of participants were engaged in multiple experimental sessions, and the experimental setting was similar to a normal classroom context.

8.2 Pedagogical implications

In view of the above theoretical and practical significances, the present research has interesting implications for second language instruction. First and foremost, teachers in task-based settings are recommended to include post-task activities in their teaching practice. The present study took transcribing as one of the candidates and showed a striking effect for improvements in formal language aspects. The procedure is perfectly feasible in regular classrooms in that only recorders and pens are needed for transcribing and the average time for transcribing a 1-min extract is around five minutes, both of which are manageable in L2 classrooms. In addition, other types of post-task activities can also be examined in further research so as to provide more focus-on-form options for pedagogical application.

In the second place, the findings highlighted the need to monitor carefully

the various conditions of post-task transcribing. Not all the transcribing conditions are beneficial for overall language improvement. For example, the pair transcribing condition is favorable for syntactic complexity improvement, while the individual condition supports an increase in lexical sophistication. For L2 learners at different proficient levels, the emphases for language development vary to a great extent. This means teachers should carefully design the transcribing conditions for different achievements in language performance.

Thirdly, teachers need to understand the factors that impact in contrasting ways on different performance aspects. For instance, the effect of revision is complex. It is generally accepted among teachers that the involvement of revision is helpful for L2 learners (Willis & Willis, 2007). However, the results reveal that revision in a general sense facilitates improvement in accuracy, but may hinder the use of complex language. Thus, we should be cautious to adopt further revision in post-task transcribing. One strategy that could be employed is to emphasize the focus of revision on both error-correction and structural improvement prior to revision. This may help learners direct their attention to both aspects. It is possible that this might reduce the negative effect on complexity to a certain extent.

Furthermore, teachers need to pay careful attention to the interaction effects between different factors in the study. The present study suggests that there is a matching effect between pair/individual transcribing and the dialogic/monologic tasks. When a transcribing condition matched the task type (either both in pairs or

both individually), participants produced longer AS-units in performance. It may be hypothesized that this kind of matching is facilitative for language complexity. Similarly, it was found that when individual revision followed individual transcribing, the effects on lexical sophistication are enhanced significantly. Put differently, the two individual conditions positively interacted with each other. Teachers are advised to be aware of those interaction effects of different elements to achieve positive outcomes in L2 teaching.

Last but not least, teachers need to be informed of the distinctive task effects on performance. Narrative tasks are popular in language assessment since they are easily handled in examination settings. However, in line with the literature (Skehan & Foster, 1997, 2001; Foster & Skehan, forthcoming), the present study reveals that interactive tasks are more promising to demonstrate treatment effects on most of the language aspects in performance. It is recommended, therefore, that interactive tasks are adopted in both L2 classrooms and assessment settings.

Finally, it should be acknowledged that all the above pedagogical recommendations which are based on the present single study cannot be warranted unless further replication studies are carried out. It should be noted as well that transcribing, when adopted as a type of post-task activity, may be beneficial for learners to focus on form, but might not necessarily bring about an immediate improvement in L2.

8.3 Limitations and further research

Given the relative infancy of post-task research, it is important to recognize the limitations of the present research. Further studies are proposed based on those limitations.

In the current study, evidence provided for language improvement was based on the quantitative analyses of the task performance. In discussing the results (c.f. chapter 7), it was noted that it was difficult to ascertain what learners did when they were transcribing, and how they oriented their attention while performing the tasks, transcribing or doing the further revision. As Ellis commented on most production studies in task-based research (2003), this research is product-oriented and has neglected process. Therefore, there is a need to extend comparable research into both quantitative and qualitative perspectives, namely to adopt a process-product approach in further studies. A series of qualitative research methods can be employed, for instance, think-aloud, stimulated recall or interview, to induce the process oriented data. The findings based on both quantitative and qualitative data could be used to provide a more comprehensive picture of post-task transcribing and make clearer the role the foreknowledge of post-task requirement plays and the effects of its operationalization. Such research could also guide more informed discussions about the benefits and the underlying processes associated with different transcribing conditions.

This study investigated the achievement of participants during a period of

four weeks. Although the time duration is relatively long as compared to previous related studies, it is still not enough to establish whether the treatment effects are durable or short-lived. Further, it is hard to judge whether the involvement of post-task transcribing leads to L2 acquisition, since the language improvement during the four weeks can only be regarded as the treatment effect on task performance rather than on L2 acquisition. As such, there is a need for longitudinal studies on the effects of task practice and post-task activities. In an EFL context like the present one, long-term effects of treatments may be investigated more readily since there is not as much extra exposure to the L2 as in ESL settings.

The current research explored the effects of post-task focus on form activities among EFL tertiary students. As non-English majors in a university, participants did not need to pay as much attention to achievement in language learning as they did in secondary schools. It is, therefore, necessary to investigate post-task activities in different settings. For example, secondary students could be the participants in future research. Further, in the present study, participants were treated as a homogeneous pool with no account of different proficiency levels. In the discussion, it was recognized that the homogeneity of participants would cause problems in identifying and correcting peer's errors, thereby generating no significant improvements among pair revision groups. In further studies, the proficiency differences can be taken into consideration so as to see whether

post-task transcribing has different effects on learners of different proficient levels, and to explore whether pair transcribing and pair revision would generate significant effects in heterogeneous pairs.

This study attempted to investigate the role of teacher revision after post-task transcribing, with no significant effect noted. One of the reasons appears due to the operationalization of the teacher revision condition. The researcher, who is not the regular teacher of the participants, acted as the teacher in the experiment. There are two disadvantages for researcher revision. On the one hand, the researcher is not aware of the participants' developmental stage and their language needs. The revision of participants' transcripts may not be as insightful as their regular teachers would achieve. On the other hand, participants were asked only to read the revised transcripts for a brief period at the beginning of next meeting which may not push participants into further noticing and processing. As a consequence, there was not any significant effect with the teacher revision condition. Given the significant role teachers have played in post-task transcribing in previous research (Lynch, 2007) and in a task-based pedagogy (Samuda, 2001; Thornby, 1997), it is argued that teacher's involvement in post-task activities is necessary. In further studies, it would be helpful to invite the regular teacher to revise the transcripts, and sufficient time needs to be organized for participants to review or process the teacher-revised version at a deeper level.

There are variables which interact in a highly complex way in this study. It is

not easy to disentangle the effects of those factors. For example, the task effects and practice effects interrelate with each other (c.f. chapter 6). Despite the fact that participants were involved in four experimental sessions, the practice effects were analysed separately for different task types. Put differently, since participants did two narrative tasks and two interactive tasks, the second narrative task performance was only compared to the first narrative so as to explore the task practice effects, and the interactive task performances were treated in the same way. The accumulated effects of task practice in the *four* task sessions were ignored. However, if performances in each of the four sessions are compared to one another, it would be hard to determine whether the improvement of the final performance is due to the advantage of a certain task type (e.g. the interactive tasks) or the involvement of task practice. This dilemma calls for further studies which disentangle the interacting variables in different ways. For example, to explore the task effects and practice effects in one study, task types can be designed as the between-subject variable and practice session as the within-subject factor. In this way, the two variables would not influence each other in the experiment.

The last suggestion for future research is not based on any limitation of the present research. The effect of individual work was found to be significant in terms of lexical sophistication. In the literature, the effect of individual work has seldom been researched. While it is reasonable for most researchers to advocate

the adoption of interactive activities, the role of the individual condition needs attention as well. Further studies are necessary to investigate the underlying differences between the pair and the individual learning conditions, as well as the effects of individual L2 learning activities on other aspects of language performance.

8.4 Conclusion

Second language acquisition is a complex phenomenon. So is focus on form research. "Researchers are torn between the desire to test theoretical claims about L2 acquisition, which requires the investigation of precise and discrete instructional options, and the desire to ensure that form-focused instruction is ecologically valid, which leads to combining options into treatments that are pedagogically defensible" (Ellis, 2008, p.900). This research has been an exploration in terms of both theoretical and pedagogical issues. In particular, the current study a) tested the proposal for the adoption of post-task activities to promote a focus on form effect; b) investigated the effects of different pedagogical conditions during post-task transcribing; and c) examined the roles of task practice and different task types. The findings have underscored the necessity for task-based research and pedagogy to give equal weight to a post-task focus on form option as they have previously to the pre- and during-task stages. As Skehan noted (2007), a task-based approach has much to offer form-focused instruction in

a variety of ways. Focus on form at the post-task stage is a promising area which is worthy of future exploration.

References

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

Anderson, A., & Lynch, T. (1988). *Listening*. Oxford: Oxford University Press.

Anderson, J.R. (1993). *Rules of the mind*. Lawrence Erlbaum Associates

Bardovi-Harlig, K. (1992). A second look at T-unit analysis: Reconsidering the sentence. *TESOL Quarterly*, 26, 390-395.

Bardovi-Harlig, K., & Bofiman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition*, 11, 17-34.

Bialystok, E. (1990). *Communication Strategies: A Psychological Analysis of Second Language Use*. Oxford: Basil Blackwell.

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

Bygate, M. (1996). Effects of task repetition: appraising the developing languages of learners. In Willis, J., & Willis, D (Eds.) (pp.136-145). Macmillan Heinemann.

Bygate, M. (2000). Introduction. *Language Teaching Research*, 4, 185-192.

Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M.Swain (Eds.) (pp.23-48). Harlow: Longman.

Bygate,M., Skehan, P., & Swain, M. (Eds.). (2001). *Researching pedagogic tasks: Second language learning, teaching and testing*. Harlow: Longman.

Byrnes, H., Maxim, H. H., & Norris, J. M. (in preparation/2010). Realizing advanced L2 writing development in a collegiate curriculum: From outcomes expectations to assessment. Monograph Series, *Modern Language Journal*, 94 (5).

Clennell, C. (1999). Promoting pragmatic awareness and spoken discourse skills with EAP classes. *ELT Journal*, 53, 83-91.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Cooper, T. C. (1976). Measuring written syntactic patterns of second language learners of German. *The Journal of Educational Research*. 69, 176-183.

Crookes, G. (1989). Planning and interlanguage variability. *Studies in Second Language Acquisition*, 11, 367-383.

Cumming, A., & Mellow, D. (1996). An investigation into the validity of written indicators of second language proficiency. In A. Cummings & R. Berwick (Eds.), *Validation in language testing* (pp.72-93). Clevedon, England: Multilingual Matters.

Daller, H. & Xue, H. (2007). Lexical richness and the oral proficiency of Chinese EFL students. In H.Daller, J. Milton & J. Treffers-Daller (Eds.) (pp.150-164). Cambridge: Cambridge University Press

Daller, H., Milton, J., & Treffers-Daller, J (Eds.). (2007). *Modelling and assessing*

vocabulary knowledge. Cambridge: Cambridge University Press.

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

Dekeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language. In C. Doughty, and J. Williams (Eds.) (pp.42-63). New York: Cambridge University Press.

Dekeyser, R. (2001). *Automaticity and automatization*. In P. Robinson, (Ed.) (pp.125-151). Cambridge: Cambridge University Press.

Dekeyser, R. (Ed.). (2007). *Practice in a second language: Perspectives from applied linguistics and cognitive psychology*. Cambridge: Cambridge University Press.

Dornyei, Z. (2002). The motivational basis of language learning tasks. In P. Robinson (Ed.), *Individual differences in second language acquisition* (pp137-158). Amsterdam: John Benjamins.

Dornyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Mahwah, NJ: Lawrence Erlbaum.

Dornyei, Z., & Kormos, J. (2000). The role of individual and social variables in oral task performance. *Language Teaching Research*, 4, 275-300.

Dornyei, Z., & Tseng, W. T. (2009). Motivational processing in interactional tasks. In A. Mackey & C. Polio (Eds.), *Multiple perspectives on interaction: Second language research in honor of Susan M. Gass* (pp.117-134).

Routledge: Taylor & Francis Group.

Doughty, C. (2001). Cognitive underpinnings of focus on form. In P. Robinson (Ed.) (pp. 206-257). Cambridge: Cambridge University Press.

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

Doughty, C., & Varela, E. (1998). Communicative focus on form. In C. Doughty & J. J. Williams (Eds.) (pp.114-138). Cambridge: Cambridge University Press.

Doughty, C., & Williams, J. (1998b). Issues and terminology. In C. Doughty. & J. Williams (Eds.) (pp.1-11). Cambridge: Cambridge University Press.

Doughty, C., & Williams, J. (1998c). Pedagogical choices in focus on form. In C. Doughty. & J. Williams (Eds.) (pp.197-261). Cambridge: Cambridge University Press.

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

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

Ellis, R. (1987). Interlanguage variability in narrative discourse: Style shifting in the use of the past tense. *Studies in Second Language Acquisition*, 9, 1-20.

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

Ellis, R. (1993). Second language acquisition and the structural syllabus. *TESOL Quarterly*, 27, 91-113.

Ellis, R. (1994). *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Ellis, R. (1997). SLA and language pedagogy: An educational perspective. *Studies in Second Language Acquisition*, 19, 69-92.

Ellis, R. (2001). Introduction: Investigating Form-Focused Instruction. *Language Learning*, 51, 1-46.

Ellis, R. (2002). The place of grammar instruction in the second/foreign language curriculum. In E. Hinkel & S. Fotos (Eds.), *New perspectives on grammar teaching in second language classrooms* (pp.17-34). Mahwah, NJ: Lawrence Erlbaum.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.

Ellis, R. (2005). *Planning and task performances in a second language*. Amsterdam: John Benjamins.

Ellis, R. (2008). *The study of second language acquisition* (2nd edition). Oxford: Oxford University Press.

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford: Oxford

University Press.

Ellis, R., Basturkmen, H., & Loewen, S. (2001). Preemptive focus on form in the ESL classroom. *TESOL Quarterly*, 35, 407-432.

Ferris, D.R. (2003). *Response to student writing: Implications for second language students*. Mahwah, NJ: Lawrence Erlbaum Associates.

Fischer, R. A. (1984). Testing written communicative competence in French. *Modern Language Journal*, 68, 13-20.

Foster, P. (1998). A classroom perspective on the negotiation of meaning. *Applied Linguistics*, 19 (1), 1-23.

Foster, P., & Skehan, P. (1996). The influence of planning on performance in task-based learning. *Studies in Second Language Acquisition*, 18, 299-324.

Foster, P., & Skehan, P. (1999). The effect of source of planning and focus on planning on task-based performance. *Language Teaching Research*, 3, 185-215.

Foster, P., & Skehan, P. (forthcoming). The effects of post-task activities on the accuracy of language during task performance. English Department, Chinese University of Hong Kong.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21, 354-375.

Fotos, S., & Nassaji, H. (Eds.). (2007). *Form-focused instruction and teacher education: Studies in honor of Rod Ellis*. Oxford University Press.

Freed, B.F. (2000). Is Fluency, Like Beauty, in the Eyes (and Ears) of the Beholder? In H. Riggenbach. (Ed.) (pp.243-265). Anna Arbor: University of Michigan Press.

Garcia Mayo, M. P. (Ed.). (2007). *Investigating tasks in formal language learning.* Multilingual Matters Ltd.

Gass, S. (1991). Grammar instruction, selective attention and learning. In R. Phillipson, E. Kellerman, L. Selinker, M. Sharwood Smith, & M. Swain (Eds.). *Foreign/Second Language Pedagogy Research.* Clevedon: Multilingual Matters.

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

Gass, S., & Varonis, M. (1985). Task variation and nonnative/nonnative negotiation of meaning. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp.149-161). Rowley, Mass: Newbury House.

Gass, S., Mackey, A., Alvarez-Torres, J. M., & Fernández-García, M. (1999). The Effects of Task Repetition on Linguistic Output. *Language Learning*, 49, 549-581.

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

Hall, J.K., & Walsh, M. (2002). Teacher-student interaction and second language learning. *Annual Review of Applied Linguistics*, 22, 186-203.

Harley, B. (1989). Functional grammar in French immersion: A classroom

experiment. *Applied Linguistics*, 10, 331-359.

Hinkel, E. (Ed). (2005). *Handbook on research in second language teaching and learning*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hinkel, E., & Fotos, S. (2002). *New Perspectives on Grammar Teaching in Second Language Classrooms*. Mahwah, NJ: Laurence Erlbaum Associates.

Hunt, K.W. (1965). *Grammatical structures written at three grade levels*. Urbana, IL: The National Council of Teachers of English.

Hyland, F. (1998). The impact of teacher written feedback on individual writers. *Journal of Second Language Writing*, 7, 255-286.

Hyland, K., & Hyland, F. (2006). *Feedback in second language writing: Contexts and Issues*. Cambridge University Press.

Iwashita, N. (2006). Syntactic complexity measures and their relation to oral proficiency in Japanese as a foreign language. *Language Assessment Quarterly*, 3, 151-169.

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

Izumi, S., Bigelow, M., Fujiwara, M., & Fearnow, S. (1999). Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition*, 21, 421-452.

Johnson, K. (1996). *Language teaching and skill learning*. Oxford: Blackwell.

Kepner, C. G. (1991). An experiment in the relationship of types of written

feedback to the development of second-language writing skills. *Modern Language Journal*, 75, 305-313.

Koponen, M., & Riggenbach, H. (2000). Overview: Varying Perspectives on Fluency. In H. Riggenbach. (Ed) (pp.5-24). Anna Arbor: University of Michigan Press.

Kormos, J. (1999). Monitoring and self-repair. *Language Learning*, 49, 303-342.

Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, N.J: Lawrence Erlbaum.

Kowal, M., & Swain, M. (1997). From semantic to syntactic processing: How can we promote it in the immersion classroom? In R.K. Johnson & m. Swain (Eds.), *Immersion education: International perspectives* (pp.284-309). Cambridge, UK: Cambridge University Press.

Krashen, S. (1982). *Principles and Practice in Second Language Acquisition*. Oxford: Pergamon.

Krashen, S. (1985). *The Input Hypothesis: Issues and implications*. London: Longman.

Lantolf, J. P. (2000). Introducing sociocultural theory. In J.P. Lantolf (Ed). *Sociocultrual theory and second language learning* (pp. 1-26). Oxford: Oxford University Press.

Lantolf, J.P. (2006). Sociocultural theory and L2: State of the Art. *Studies in Second Language Acquisition*, 28, 67-109.

Laufer, B. (1994). The lexical profile of second language writing: Does it change over time? *RELC Journal*, 25, 21-33.

Lee, G., & Schallert, D. L. (2008). Meeting in the margins: Effects of the teacher-student relationship on revision processes of EFL college students tasking a composition course. *Journal of Second Language Writing*, 17, 165-182.

Lee, I. (2004). Error correction in L2 secondary writing classrooms: The case of Hong Kong. *Journal of Second Language Writing*, 13, 285-312.

Leeman, J.(2007). Feedback in L2 learning: Responding to errors during practice. In R. Dekeyser (Ed.) (pp.111-137). Cambridge: Cambridge University Press.

Leeser, M. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research*, 8, 55-81.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40, 387-417.

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach. (Ed) (pp.25-42). Anna Arbor: University of Michigan Press.

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

Lightbown, P. (2000). Classroom SLA research and second language teaching, *Applied Linguistics*, 21, 431-462.

Lightbown, P., & Spada, N. (1990). Focus-on-form and corrective feedback in communicative language teaching: effects on second language learning.

Studies in Second Language Acquisition, 12, 429-448.

Loewen, S. (2005). Incidental focus on form and second language learning. *Studies in Second Language Acquisition*, 27, 361-386.

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

Long, M. (1985). A role of instruction in second language acquisition: task-based language teaching. In K. Hyltenstam & M. Peineman (Eds.). *Modelling and assessing second language acquisition* (pp.77-99). Clevedon: Multilingual Matters.

Long, M. (1988). Instructed interlanguage development. In L. Beebe, (Ed.), I*ssues in Second Language Acquisition: Multiple Perspective* (pp.115-141). Rowley, Mass: Newbury House.

Long, M. (1991). Focus on form: A design feature in language teaching methodology. In K.de Bot, R, Ginsberg, & C. Kramsch (Eds.), *Foreign language research in cross-cultural perspective* (pp.39-52). Amsterdam: John Benjamins.

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

Long, M., & Crookes, G. (1992). Three approaches to task-based syllabus design. *TESOL Quarterly*, 26, 27-56.

Long, M., & Robinson, P. (1998). Focus on form: Theory, research and practice.

In C. Doughty and J. Williams (Eds.) (pp.15-41). Cambridge: Cambridge University Press.

Lynch, T. (2001). Seeing what they meant: Transcribing as a route to noticing. *ELT Journal*, 55, 124-132.

Lynch, T. (2007). Learning from the transcripts of an oral communication task. *ELT Journal*, 61(4), 311-320.

Lynch, T., & Maclean, J. (2001). A case of exercising: Effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan, & M.Swain (Eds.) (pp.141-162). Harlow: Longman.

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

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

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

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

MacWhinney, B. (1995). Language–specific prediction in foreign language learning. *Language Testing*, 12, 292-320.

MacWhinney, B. (1997). Implicit and explicit processes: Commentary. *Studies in Second Language Acquisition*, 19, 227-282.

Malvern, D., & Richards, B. (1997). A new measure of lexical diversity. In A. Ryan. and A. Wray (Eds.), *Evolving models of language: Papers from the annual meeting of the British Association for Applied Linguistics* held at the University of Wales, Swansea, September, 1996. Clevedon, UK: Multilingual Matters, 58-71.

Malvern, D., & Richards, B. (2002). Investigating accommodation in language proficiency interviews using a new measure of lexical diversity. *Language Testing*, 19, 85-104.

Mckee, G., Malvern, D., & Richards, B. (2000). Measuring vocabulary diversity using dedicated software. *Literary and Linguistic Computing*, 15, 323-338.

McLaughlin, B. (1987). *Theories of second language learning*. London: Arnold.

McLaughlin, B. (1990). Restructuring. *Applied Linguistics*, 11, 113-128.

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: an information-processing perspective. *Language Learning*, 33, 135-158.

Meara, P., & Bell, H. (2001). P_Lex: A simple and effective way of describing the lexical characteristics of short L2 text. *Prospect*, 16, 5-19.

Mehnert, U. (1998). The effects of different lengths of time for planning on second language performance. *Studies in Second Language Acquisition*, 20, 83-108.

Mennim, P. (2003). Rehearsed oral L2 output and reactive focus on form. *ELT Journal*, 57, 130-138.

Mochizuki, N., & Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research*, 12, 11-37.

Montgomery, J. L., & Baker, W. (2007). Teacher-written feedback: Students perceptions, teacher self-assessment, and actual teacher performance. *Journal of Second Language Writing*, 16, 82-99.

Muranoi, H. (2007). Output practice in the L2 classroom. In R. Dekeyser (Ed.) (pp.51-82). Cambridge: Cambridge University Press.

Musumeci, D. (1996). Teacher-learner negotiation in content-based instruction: Communication at cross communicative task in EFL classrooms. *Language Learning*, 50, 617-673.

Nassaji, H. (2000). Towards integrating form-focused instruction and communicative interaction in the second language classroom: Some pedagogical possibilities. *The Modern Language Journal*, 84, 241-250.

Nassaji, H., & Fotos, S. (2004). Current developments in the teaching of grammar. *Annual Review of Applied Linguistics*, 24, 126-145.

Nation, I. S. P. (2007). Fundamental issues in modelling and assessing vocabulary knowledge. In H.Daller, J. Milton & J. Treffers-Daller (Eds.) (pp.35-43). Cambridge: Cambridge University Press

Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second

language acquisition. *ELT Journal*, 47, 203-210.

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

Norris, J., & Ortega, L. (2008). Measurement for understanding: The case of complexity, accuracy and fluency. AAAL 2008, Washington DC.

Nunan, D. (1993). Task-based Syllabus Design: Selecting, Grading and Sequencing Tasks. In G. Crookes and S. Gass. (Eds). *Tasks in a pedagogical context : Integrating theory and practice* (pp.55-68). Clevedon, UK: Multilingual Matters.

O'Donnell, R. C. (1976). A critique of some indices of syntactic maturity. *Research in the Teaching of English*, 10, 33-38.

Ortega, L. (1999). Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21, 108-148.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 24, 492-518.

Ortega, L. (2007). Meaningful L2 practice in foreign language classrooms: A cognitive-interactionist SLA perspective. In R. Dekeyser (Ed.) (pp180-207). Cambridge: Cambridge University Press

Pica, T. (1994). Research on negotiation: What does it reveal about second language acquisition? *Language Learning*, 44, 493-527.

Pica, T. (1997). Second language teaching and research relationships: A North American view. *Language Teaching Research*, 1, 48-72.

Pica, T. (2002). Subject-matter content: How does it assist the interactional and linguistic needs of classroom language learners? *Modern Language Journal*, 86, 1-19.

Pica, T., Holliday, L., Lweis, N., & Moregenthaler, L. (1989). Choosing and using communicative tasks for second language research and instruction. In S. Gass and G. Crookes (Eds.), *Tasks and language learning: Integrating theory and practice* (pp.9-34). Clevedon, UK: Multiginlugal Matters

Polio, C. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning, 47*, 101-143.

Polio, C., & Gass, S. (1998). The role of interaction in narrative speaker comprehension of non-native speakers speech. *Modern Language Journal*, 82, 308-319.

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford, UK: Oxford University Press.

Read, J. (2000). *Assessing vocabulary*. Cambridge: CUP.

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

Reber, A. (1993). *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*. Oxford: Oxford University Press.

Richards, B. J., & Malvern, D. (2007). Validity and threats to the validity of vocabulary measurement. In H.Daller, J. Milton & J. Treffers-Daller (Eds.) (pp.79-92). Cambridge: Cambridge University Press

Riggenbach, H. (2000). *Perspective on Fluency*. Anna Arbor: University of Michigan Press.

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

Robinson, P. (2001a). *Cognition and second language instruction*. Cambridge: Cambridge University Press.

Robinson, P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22, 27-57.

Russell, J., & Spada, N. (2006). The effectiveness of corrective feedback for the acquisition of L2 grammar: A meta-analysis of the research. In J.M. Norris & L. Ortega (Eds.), *Synthesizing research on language learning and teaching* (pp.133-162). John Benjamins Publishing Company.

Rutherford, W. (1988). *Second language grammar: Teaching and learning*. London: Longman.

Salomon, G., & Globerson, T. (1989). When teams do not function the way they ought to. *International Journal of Education Research*, 13, 89-98.

Samuda, V. (2001). Guiding relationships between form and meaning during task performance: The role of the teacher. In M. Bygate, P. Skehan, & M.Swain

(Eds.) (pp.119-140). Harlow: Longman.

Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. Palgrave Macmillan

Sanguram, J. (2005). The effects of focusing on meaning and form in strategic planning. In R. Ellis (Ed.) (pp.111-142). Amsterdam: John Benjamins.

Savignon, S. (2005). Communicative language teaching: Goals and Strategies. In E. Hinkel (Ed.) (pp.635-652). Mahwah, NJ: Lawrence Erlbaum Associates.

Schachter, J. (1988). Second language acquisition and its relationship to Universal Grammar. *Applied Linguistics*, 9, 219-235.

Schenkein, J. (1980). A taxonomy for repeating action sequences in natural conversation. In B. Butterworth (Ed.), *Language Production*, Vol 1. Speech and talk (pp.21-47). New York: Academic Press.

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

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*. 13, 206-226.

Schmidt, R. (2001). Attention. In P. Robinson. (Ed.) (pp.3-28). Cambridge: Cambridge University Press.

Schmidt, R., & Frota, S. (1986). Developing basic conversational ability in a second language: A case-study of an adult learner. In R. Day (Ed.), *Talking to learn: conversation in second language acquisition* (pp.237-322). Rowley,

Mass: Newbury House.

Shehadeh, A. (2001). Self and other imitated modified output during task-based interaction. *TESOL Quarterly*, 35, 433-457.

Sheppard, K. (1992). Two feedback types: Do they make a difference? *RELC Journal*, 23, 103-110.

Skehan, P. (1996a). A framework for the implementation of task-based instruction. *Applied Linguistics*, 37, 38-62.

Skehan, P. (1996b). Second language acquisition research and task-based instruction. In J. Willis and D. Willis (Eds.) (pp.17-30). Macmillan Heinemann.

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

Skehan, P. (2001). Tasks and language performance. In M. Bygate, P. Skehan, & M. Swain (Eds.). Harlow: Longman.

Skehan, P. (2003). Task-based instruction. *Language Teaching*. 36, 1-14.

Skehan, P. (2007). Task research and language teaching: Reciprocal relationships. In Fotos, S., & Nassaji, H (Eds.) (pp.55-69). Oxford University Press.

Skehan, P. (2009). Lexical performance by native and Non-native speakers on language-learning tasks. In B. Richards, H. Daller, D.D. Malvern, P. Meara, J Milton, B, Richards & J. T. Daller (Eds.), *Vocabulary studies in first and second language acquisition: The interface between theory and application*

(pp.107-124). London: Palgrave Macmillan.

Skehan, P., & Foster, P. (1997). The influence of planning and post-task activities on accuracy and complexity in task based learning. *Language Teaching Research*, 1, 185-211.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retelling. *Language learning*, 49, 93-120.

Skehan, P., & Foster, P. (2001). Cognition and tasks. In P.Robinson (Ed.) (pp.183-205). New York: Cambridge University Press.

Skehan, P., & Foster, P. (2005). Strategic and on-line planning: The influence of surprise information and task time on second language performance. In R. Ellis (Ed.) (pp.193-216). Amsterdam: John Benjamins.

Skehan, P. (2005). coding manual for task profile program. The Chinese University of Hong Kong.

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

Spada, N. (2005). Conditions and challenges in developing school-based SLA research programs. *The Modern Language Journal*, 89, 328-338.

Stern, H. (1983). *Fundamental concepts of language teaching*. Oxford: Oxford University Press.

Storch, N. (2005).Collaborative writing: product, process, and students' reflections. *Journal of Second Language Writing*, 14, 153-173.

Storch, N. (2007). Investigating the merits of pair work on a text editing task in ELS classes. *Language Teaching Research*, 11, 143-159.

Storch, N., & Wigglesworth, G. (2007). Writing tasks: The effects of collaboration. In M.P. Garcia Mayo (Ed.) (pp.157-177). Multilingual Matters Ltd.

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

Swain, M. (1988). Manipulating and complementing content teaching to maximize second language learning. *TESOL Canadian Journal*, 6, 68-83.

Swain, M. (1991). Manipulating and complementing content teaching to maximize second language learning. In E.Kellerman, R. Phillipson, L. Selinker, M. Sharwood Smith, & M. Swain (Eds.), *Foreign/second language pedagogical research* (pp.234-250). Clevedon: Multilingual Matters.

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

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidhofer., (Eds.), *Principles and practice in applied linguistics: Studies in honor of H.G. Widdowson* (pp 125-144).Oxford, UK: Oxford University Press.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.) (pp.64-82). Cambridge: Cambridge University Press.

Swain, M. (2005). The output hypothesis: theory and research. In E. Hinkel (Ed.) (pp.471-483). Mahwah, NJ: Lawrence Erlbaum Associates.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391.

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: Two adolescent French immersion students working together. *Applied Linguistics*, 16, 371-391.

Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan, & M. Swain (Eds.). Harlow: Longman.

Tarone, E. (1985). Variability in interlanguage use: a study of style-shifting in morphology and syntax. *Language Learning*, 35, 373-403.

Tavakoli, P. & Skehan, P. (2005). Planning, task structure, and performance testing. In R. Ellis (Ed) (pp.239-271). Amsterdam: John Benjamins

Thornbury, S.(1997). Reformulation and reconstruction: Tasks that promote 'noticing'. *English Language Teaching Journal*, 51, 326-335.

Toth, P. (2008). Teacher-led and learner-led discourse in task-based grammar instruction: Providing procedural assistance for L2 morphosyntactic development. *Language Learning*, 58, 237-283.

Truscott, J. (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327-369.

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

Truscott, J. (2007). The effect of error correction on learners' ability to write accurately. *Journal of Second Language Writing*, 16, 255-272.

Van den Branden, K. (1997). Effects of negotiation on language learners' output. *Language Learning*, 47, 589-636.

Van den Branden, K. (2007). Second language education: Practice in perfect learning conditions? In R. Dekeyser (Ed.) (pp.161-179). Cambridge: Cambridge University Press.

VanPatten, B. (1990). Attention to form and content in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, 12, 287-301.

VanPatten, B. (1994). Evaluating the role of consciousness in SLA: terms, linguistic features, and research methodology. *AILA Review*, 11, 27-36.

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. CT:Ablex.

VanPatten, B. (2002). Processing instruction: an update. *Language Learning*, 52, 755-804.

VanPatten, B. (2004). *Processing instruction: Theory, research and commentary*. Chicago Lawrence Erlbaum Associates Publishers.

Varonis, E. & Gass, S. (1985). Non-native/non-native conversation: A model for negotiation of meaning. *Applied Linguistics*, 6, 71-90.

Vocate, D. R. (1994). Self-talk and inner speech: Understanding the uniquely human aspects of intrapersonal communication. In D. R. Vocate (Ed). *Intrapersonal communication: Different voices, different minds* (pp.3-31). Mahwah, NJ: Erlbaum.

Vygotsky, L.S. (1978). *Mind in society. The development of higher psychological processes*. Cambridge, Mass: Harvard University Press.

Wang, Zhan. 2009. *Modeling L2 Speech Production and Performance: Evidence from Five Types of Planning and Two Task Structures*. Unpublished Ph.D.Thesis. The Chinese University of Hong Kong.

Wertsch, J. V. (1991). *Voices of the mind: A sociocultural approach to mediated action*. Cambridge, Mass: Harvard University Press.

Widdowson, H. (1989). Knowledge of language and ability for use. *Applied Linguistics*, 10, 128-137.

Wigglesworth, G. (1997). An investigation of planning time and proficiency level on oral test discourse. *Language Testing*, 14, 85-106.

Wilkins, D. (1976). *Notional syllabuses*. Oxford: Oxford University Press.

Williams, J. (1999). Learner-generated attention to form. *Language Learning*, 49, 583-625.

Williams, J. (2005). Form-focused instruction. In E. Hinkel. (Ed.) (pp.671-692). Mahwah, NJ: Lawrence Erlbaum Associates.

Willis, D., & Willis, J. (2007). *Doing task-based teaching*. Oxford: Oxford University Press.

Willis, J. (1996). *A framework for task-based learning*. Addison Wesley Longman Limited.

Willis, J., & Willis, D. (Eds.). (1996). *Challenge and change in language teaching*. Macmillan Heinemann.

Wolfe-Quintero, K., Inagaki, S., & Kim, H.Y. (1998). *Second language development in writing: Measures of fluency, accuracy and complexity* (Technical Report #17). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Center.

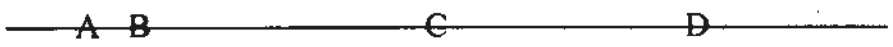
Yau, M. (1991). The role of language factors in second language writing. In L. Malave & G. Duquette (Eds.), *Language, culture and cognition: A collection of studies in first and second language acquisition* (pp.266-283). Clevedon, England: Multilingual Matters.

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

Appendix I Rating Scheme for Video Selection

Four tertiary teachers were asked to rate the four episodes in terms of required background knowledge, understanding difficulty, retelling difficulty as follows. A, B, C, and D represent the four video episodes which they have watched before marking. 1, 2, 3, 4 and 5 indicate the degree of difficulty regarding different aspects.

Example for marking:



1 2 3 4 5
easy difficult

a. As far as the background knowledge is concerned, please mark the difficulty level in the following rating scale:

1 2 3 4 5
easy difficult

b. As far as the understanding difficulty is concerned, please mark the difficulty level:

1 2 3 4 5
easy difficult

c. Suppose you are to retell the story. As far as the vocabulary is concerned, please mark the difficulty level:

1 2 3 4 5
easy difficult

d. Suppose you are to retell the story. As far as the grammar is concerned, please mark the difficulty level:

1 2 3 4 5
easy difficult

Appendix II Topics and materials for discussion tasks in Pilot I

Task introduction for the first topic "Scholarship Award":

You are the student members in a selective committee of a university. Among the two student candidates, one should be selected as qualified to get the outstanding student scholarship. The profiles of the candidates are listed in the following:

Student A: Maria is a senior student in this university. She is excellent in her academic work. She has won various prizes in academic competitions. It is worthy of mentioning that she won the first prize in an international academic competition this year. However, some of her classmates think she is not active enough that she seldom joins the class activities for sports and entertainment. As such, she has few friends in the class.

Student B: Bob is a fresh student in this university. He is the monitor of his class. He is the chief organizer and director of a project named "Hand in Hand with a Poor Kid". All his classmates are invited to visit some poor kids in an undeveloped mountainous area. Then, every student keeps in touch with a poor kid so as to support the kid to finish his/her elementary education. This project enjoys a well-known reputation in this city. However, Bob's academic achievement is only at an average level.

Task Instruction for Jury Discussion :

It is adapted from a news story. After school, three boys went to the zoo which was closed after the school time. The boys climbed to the crocodile pond and used some branches to play with the crocodile. Suddenly, the crocodile got irritated and attacked the boys. One of the boys was bitten to death. The zoo was charged of being guilty of neglect, and the parents of the dead boy thought the zoo should take all the responsibility for the death of the boy. As jury members, you are supposed to decide whether the charge is reasonable or not.

Appendix III Task Instruction for Decision-making Tasks in the Main Study

In this task, you have some letters which were sent to the Problem Page of a magazine, to a "problem aunt" called Sue. She replied to each of them with advice to help writer the writer solve each of the problems.

Imagine that the two of you together are "Sue" and that your task is to agree on the advice to put in the letter you send to each of these people. In each case, to think of the different sorts of advice which are possible, of why one bit of advice would work better, or of why some bit of advice might contain difficulties or dangers. Work out what the best advice is that you could put in your letter of reply.

在很多现代杂志或报刊中，均设有类似"知心姐姐""排忧解难"的栏目。这类栏目经常收到读者来信，信中讲述自己的烦恼，希望栏目专家能帮助读者解决困难，提供合适的建议。

假设你们两人是某杂志社"知心姐姐"栏目的编辑知心姐姐（Sue）。你们现收到如下读者来信。请你们首先讨论针对该问题的各种可能的解决方法，然后逐一分析各种方法的优势和不足（例如，为什么办法 A 好过 B；若实施办法 C 有哪些危险和困难等），最后得出一个你们认为最好的解决途径。

Appendix IV Cloze Test

Passage 1:

The task of being accepted and enrolled (招收) in a university begins early for some students. Long __1__ they graduate from high school. These students take special __2__ to prepare for advanced study. They may also take one of more examinations that test how __3__ prepared they are for the university. In the final year of high school, they __4__ applications and send them, with their student records, to the universities which they hope to __5__. Some high school students may be __6__ to have an interview with representatives of the university. Neatly, __7__, and usually very frightened, they are __8__ to show that they have a good attitude and the __9__ to succeed. When the new students are finally __10__, there may be one more step they have to __11__ before registering for classes and __12__ to work. Many colleges and universities __13__ an orientation (情况介绍) program for new students. __14__ these programs, the young people get to know the __15__ for registration and student advising, university rules, the __16__ of the library and all the other __17__ services of the college or university. Beginning a new life in a new place can be very __18__. The more knowledge students have __19__ the school, the easier it will be for them to __20__ to the new environment. However, it takes time to get used to college life.

1. A) as	B) after	C) since	D) before
2. A) courses	B) disciplines	C) majors	D) subjects
3. A) deeply	B) widely	C) well	D) much
4. A) fulfill	B) finish	C) complete	D) accomplish
5. A) attend	B) participate	C) study	D) belong
6. A) acquired	B) considered	C) ordered	D) required
7. A) decorated	B) dressed	C) coated	D) worn
8. A) decided	B) intended	C) sealed	D) determined
9. A) power	B) ability	C) possibility	D) quality
10. A) adopted	B) accepted	C) received	D) permitted
11. A) make	B) undergo	C) take	D) pass
12. A) getting	B) putting	C) falling	D) sitting
13. A) offer	B) afford	C) grant	D) supply
14. A) For	B) Among	C) In	D) On
15. A) processes	B) procedures	C) projects	D) provisions
16. A) application	B) usage	C) use	D) utility
17. A) major	B) prominent	C) key	D) great
18. A) amusing	B) misleading	C) alarming	D) confusing
19. A) before	B) about	C) on	D) at
20. A) fit	B) suit	C) yield	D) adapt

Passage II:

Most children with healthy appetites (胃口) are ready to eat almost anything that is offered them and a child rarely dislikes food __1__ it is badly cooked. The __2__ a meal is cooked and served is most important and an __3__ served meal will often improve a child's appetite. Never ask a child __4__ he likes or dislikes a food and never __5__ likes and dislikes in front of him or allow __6__ else to do so. If the father says he hates fat meat or the mother __7__ vegetables in the child's hearing he is __8__ to copy this procedure. Take it __9__ granted that he likes everything and he probably __10__. Nothing healthful should be omitted from the meal because of a __11__ dislike. At meal times it is a good __12__ to give a child a small portion and let him __13__ back for a second helping rather than give him as __14__ as he is likely to eat all at once. Do not talk too much to the child __15__ meal times, but let him get on with his food, and do not __16__ him to leave the table immediately after a meal or he will __17__ learn to swallow his food __18__ he can hurry back to his toys. Under __19__ circumstances must a child be coaxed(哄骗) __20__ forced to eat.

1. A) if	B) until	C) that	D) unless
2. A) procedure	B) process	C) way	D) method
3. A) adequately	B) attractively	C) urgently	D) eagerly
4. A) whether	B) what	C) that	D) which
5. A) remark	B) tell	C) discuss	D) argue
6. A) everybody	B) anybody	C) somebody	D) nobody
7. A) opposes	B) denies	C) refuses	D) offends
8. A) willing	B) possible	C) obliged	D) likely
9. A) with	B) as	C) over	D) for
10. A) should	B) may	C) will	D) must
11. A) supposed	B) proved	C) considered	D) realized
12. A) point	B) custom	C) idea	D) plan
13. A) ask	B) come	C) return	D) take
14. A) much	B) little	C) few	D) many
15. A) on	B) over	C) by	D) during
16. A) agree	B) allow	C) force	D) persuade
17. A) hurriedly	B) soon	C) fast	D) slowly
18. A) so	B) until	C) lest	D) although
19. A) some	B) any	C) such	D) no
20. A) or	B) nor	C) but	D) neither

Passage III:

During recent years we have heard much about "race": how this race does certain things and that race believes certain things and so on. Yet, the __1__ phenomenon of race consists of a few surface indications. We judge race usually __2__ the coloring of the skin: a white race, a brown race, a yellow race and a black race. But __3__ you were to remove the skin you could not __4__ anything about the race to which the individual belonged. There is __5__ physical structure, the brain or the internal organs(器官) to __6__ a difference. There are four types of blood. __7__ types are found in every race, and no type is distinct to any race. Human brains are the __8__. No scientists could examine a brain and tell you the race to which the individual belonged. Brains will __9__ in size, but this occurs within every race. __10__ does size have anything to do with intelligence. The largest brain __11__ examined belonged to a person of weak __12__. On the other hand, some of our most distinguished people have had __13__ brains. Mental tests which are reasonably __14__ show no differences in intelligence between races. High and low test results both can be recorded by different members of any race. __15__ equal educational advantages, there will be no difference in average standings, either on account of race or geographical location. Individuals of every race __16__ civilization to go backward or forward. Training and education can change the response of a group of people, __17__ enable them to behave in a __18__ way. The behavior and ideas of people change according to circumstances, but they can always go back or go on to something new __19__ is better and higher than anything __20__ the past.

1. A) complete	B) full	C) total	D) whole
2. A) in	B) from	C) at	D) on
3. A) since	B) if	C) as	D) while
4. A) speak	B) talk	C) tell	D) mention
5. A) something	B) everything	C) nothing	D) anything
6. A) display	B) indicate	C) demonstrate	D) appear
7. A) All	B) Most	C) No	D) Some
8. A) same	B) identical	C) similar	D) alike
9. A) remain	B) increase	C) decrease	D) vary
10. A) Only	B) Or	C) Nor	D) So
11. A) ever	B) then	C) never	D) once
12. A) health	B) body	C) mind	D) thought
13. A) big	B) small	C) minor	D) major
14. A) true	B) exact	C) certain	D) accurate
15. A) Provided	B) Concerning	C) Given	D) Following
16. A) make	B) cause	C) move	D) turn
17. A) and	B) but	C) though	D) so
18. A) ordinary	B) peculiar	C) usual	D) common
19. A) that	B) what	C) whichever	D) whatever
20. A) for	B) to	C) within	D) in